(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 November 2001 (15.11.2001)

PCT

(10) International Publication Number
WO 01/85786 A2

(51) International Patent Classification[7]: C07K 14/705

(21) International Application Number: PCT/US01/14527

(22) International Filing Date: 4 May 2001 (04.05.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/202,151	5 May 2000 (05.05.2000)	US
60/227,567	23 August 2000 (23.08.2000)	US
60/247,855	13 November 2000 (13.11.2000)	US

(71) Applicant *(for all designated States except US)*: **AMERICAN HOME PRODUCTS CORPORATION** [US/US]; Five Giralda Farms, Madison, NJ 07940 (US).

(72) Inventors; and

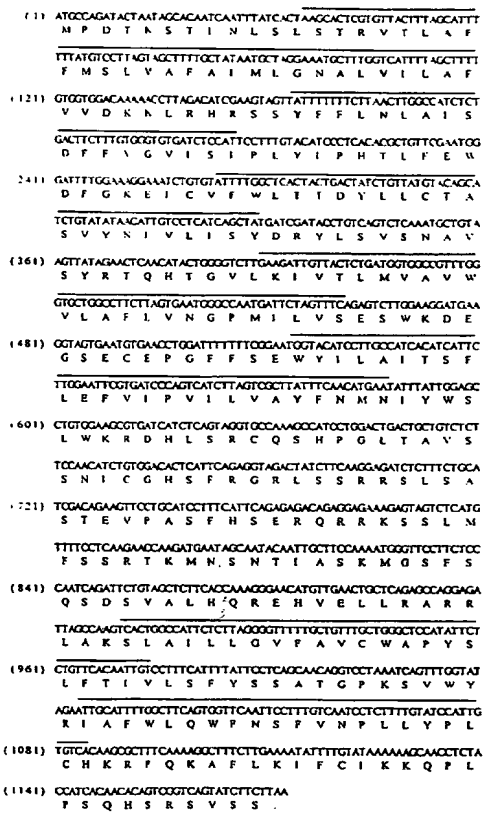
(75) Inventors/Applicants *(for US only)*: **JONES, Philip, G.** [GB/US]; 32 Rocky Brook Road, Cranbury, NJ 08512 (US). **BLATCHER, Maria** [GB/US]; 720 Cox Road, Morristown, NJ 08057 (US). **WU, Shujian** [CN/US]; 972 Ithan Lane, Langhorne, PA 19047 (US). **PAUSCH, Mark, H.** [US/US]; 23 Montgomery Avenue, Rocky Hill, NJ 08553 (US).

(74) Agents: **CHOKSI, Neepa** et al.; Darby & Darby P.C., 805 Third Avenue, New York, NY 10022-7513 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

[Continued on next page]

(54) Title: HUMAN HISTAMINE H_4 RECEPTOR



(57) Abstract: The present invention discloses the identification of a novel histamine receptor, termed H_4. Amino acid sequences, nucleic acid sequences, vectors, and host cells are also discussed. Additionally, methods of detecting agonists and antagonists for the receptor are disclosed herein.

WO 01/85786 A2

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

HUMAN HISTAMINE H_4 RECEPTOR

This application claims priority under 35 U.S.C. § 119 from provisional patent application Serial Nos. 60/202,151, filed May 5, 2000, 60/227,567, filed August 23, 2000, and 60/247,855, filed November 13, 2000; which all are hereby incorporated by reference in their entireties.

FIELD OF THE INVENTION

The present invention discloses the identification of a novel histamine receptor, termed H_4.

BACKGROUND OF THE INVENTION

Histamine produces numerous physiological effects in the body through interaction with one of three different cell surface receptors, classified as H_1, H_2, and H_3. These receptors belong to the guanine nucleotide binding protein coupled receptor (G-protein coupled receptors, GPCR) class.

Stimulation of histamine H_1 receptors produces symptoms that are typically associated with physiological responses to allergic stimuli (Ash and Schild, Br. J. Pharmacol. 1966, 27:427). These effects are blocked by H_1 antagonists such as, for example, diphenhydramine. H_1 antagonists are generally defined as "classical antihistamines". Classical antihistamines are the active ingredient in most over-the-counter allergy medications. Pharmacological studies indicate that agonist activation of these receptors stimulates the inositol phosphate pathway, and thus stimulates formation of inositol triphosphate (IP_3) and diacylglycerol (DAG).

Histamine H_2 receptors have been shown to play a role in gastric acid secretions (Black *et al.*, Nature 1972, 236:385). Histamine H_2 receptor antagonists such as, for example, cimetidine and ranitidine, are often the active ingredient in over-the-counter and prescription drugs that are used to treat duodenal ulcers, gastric ulcers, heartburn, indigestion, and other disorders of the gastrointestinal tract. Activation of histamine H_2 stimulates adenylyl cyclase activity and stimulates formation of cAMP.

Histamine H_3 receptors are a relatively new member of the histamine receptor class. These receptors were originally described as histamine autoreceptors. These receptors were shown to be located on presynaptic histamine nerve terminals and to control the rate of histamine synthesis and release in the brain (Arrang *et al.*, Nature 1983, 302:832). Recent studies indicate that H_3 receptors also are present on non-histamine presynaptic nerve terminals. H_3 receptors may modulate the synthesis and release of other neurotransmitters such as dopamine, serotonin, acetylcholine, and norepinephrine. H_3 receptors also have been found in peripheral tissues. Histamine activation of H_3 receptor inhibits adenylyl cyclase activity and formation of cAMP.

Multiple pharmacological studies have indicated the presence of H_3 receptor subtypes (Leurs *et al.* JPET 1996, 276:1009-1015; Cumming and Gjede, Brain Res. 1994, 641:203-207; Calpham and Kilpatrick, Br. J. Pharmacol 1996, 107:919-923; Schworer *et al.* Naunyn-Schmiedeberg's Arch. Pharmacol. 1994, 350:375-379; Schkicker *et al.*, Naunyn-Schmiedeberg's Arch. Pharmacol. 1996, 353:482-488). Additionally, pharmacological characterization of a histamine receptor on eosinophils describes the greater potency of histamine compared to R-α-methylhistamine (Raible, *et al.* Am. J. Respir. Crit. Care.Med 1994, 149:1506-1511.). However, the existence of these receptor subtypes has yet to be substantiated by molecular biological techniques.

SUMMARY OF THE INVENTION

The present invention contemplates an isolated histamine H_4 receptor protein having an amino acid sequence at least 51% identical or comprising at least 10 contiguous amino acids from the sequence depicted in SEQ ID NO:2. The H_4 receptor protein binds ligands comprising an imidazole and an amine, which imidazole and amine are attached by an alkyl chain, where the rank order of efficacy of modulation of second messenger formation of the ligands at the H_4 receptor protein is 5>6=10>8=4, where the number represents the number of carbons in the alkyl chain. In one embodiment, upon binding histamine or a histamine agonist the receptor protein inhibits second messenger formation. Preferably the second messenger is cAMP.

The present application also discloses an isolated H_4 receptor protein having an amino acid sequence with at least 95% sequence identity to human H_4 receptor protein having

an amino acid sequence as depicted in SEQ ID NO: 2. In a specific embodiment, the protein is a human H_4 receptor protein, *e.g.*, having an amino acid sequence as depicted in SEQ ID NO:2 or an allelic variant of that sequence.

Also provided are nucleic acids, *e.g.*, cDNAs encoding the H_4 receptor proteins as discussed above. In another embodiment, an isolated nucleic acid encoding an H_4 receptor protein, which nucleic acid hybridizes under stringent conditions to a nucleic acid having a sequence of at least 20 nucleotides identical to a corresponding nucleotide sequence of the same number of bases in SEQ ID NO:1 or its complement.

The isolated nucleic acids encoding H_4 receptors can be part of vectors, *e.g.*, for cloning, expression, and/or expansion. An expression vector comprises the nucleic acid encoding the H_4 receptor protein operably associated with an expression control sequence. The invention further provides host cells and non-human transgenic animals containing such an expressible vector, and methods for producing an H_4 receptor polypeptide using such host cells.

In addition, the invention provides an isolated nucleic acid, such as a primer or probe, of at least 10 bases having a nucleotide sequence identical to a corresponding nucleotide sequence of the same number of bases in SEQ ID NO:1, or its complement. The invention also provides an antibody that specifically binds an H_4 receptor.

The present invention further contemplates a method for detecting expression of H_4 receptor, which method comprises detecting mRNA encoding H_4 receptor in a sample from a cell suspected of expressing H_4 receptor or detecting the H_4 receptor protein with an antibody of the invention.

The present invention also contemplates an assay system for identifying H_4 receptor ligands. The assay system comprises a sufficient number of transformed host cells to be able to detect an alteration in second messenger accumulation. Preferably, the second messenger is cAMP.

The present invention also contemplates a method for identifying a test compound that antagonizes or agonizes histamine H_4 receptors. The method comprises detecting an alteration in the level of a second messenger in the assay system contacted with the test compound. In the method, an increase in the level of the second messenger indicates that the test compound antagonizes the H_4 receptor. A decrease in the level of the second messenger indicates that the test compound agonizes the H_4 receptor.

The present invention also discloses an isolated nucleic acid that specifically hybridizes under highly stringent conditions to the complement of the sequence depicted in SEQ ID NO:1, where the nucleic acid encodes a histamine H_4 receptor protein.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1. Sequence of the human H_4 cDNA (SEQ ID NO: 1) and the deduced amino acid sequence of the protein it encodes (SEQ ID NO:2). The TM domains are denoted by lines.

Figures 2A, 2B, and 2C. (A) Agonist assay, Square: histamine; triangle: R-α-methylhistamine; inverted triangle: clobenpropit; diamond: thioperamide. **(B)** Antagonist assay. Assay were conducted in duplicate and presented as the average ± SEM. **(C)** Activity in mammalian cells.

Figure 3. Antagonist activity of a series of histamine analogs. Square: clobenpropit; triangle: thioperamide. Assay were conducted in duplicate and presented as the average ± SEM.

Figure 4. H_4 expression in from 6 human T cell clones derived from a single human donor. Three of these clones were CD4+ cell clones (RG4.3B, RG4.3A, and RG4.3), two were CD8+ cell clones (RG8.1C and RG8.1A), and one was an NKT cell clone (RG1).

Figure 5. H_4 expression in bulk populations of CD4+ and CD8+ T cells isolated from three different individuals. B-CD8 and B-CD4 were CD8+ and CD4+ cells, respectively, obtained from individual 1. L-CD8 and L-CD4 were CD8+ and CD4+ cells, respectively, obtained from individual 2. U-CD8 and U-CD4 were CD8+ and CD4+ cells, respectively, obtained from individual 3.

DETAILED DESCRIPTION

The present invention is based, in part, on discovery of a novel histamine receptor, which has been termed H_4. The new histamine receptor was cloned from a human heart library using H_3 specific primers, but proved to be a new histamine receptor. Yeast and human endothelial kidney (HEK) cells were transformed with a human H_4 expression vector. Transformed cells in multi-well plates were treated with test compounds, and regulation of the intracellular second messenger cyclic adenosine monophosphate (cAMP) formation was

determined. Modulation of cAMP formation is ligand and concentration dependent.

The nucleic acid and protein sequences of H_4 shows homology to known G-protein coupled receptors. Specifically, the H_4 protein shows homology to biogenic-amine G-protein coupled receptors. As in other biogenic amine receptors, the present protein contains conserved aspartate residues within transmembrane (TM) domains 2 and 3 (at positions 61 and 94). A DY motif (D94 and Y95) in TM3 is found only in histamine and muscarinic receptors. A DRY sequence motif that exists at the cytoplasmic interface of TM3 of all GPCRs also can be found in the present sequence (amino acids 111- 113). Also present are two conserved cysteine residues in the first and second extracellular loops (position 88 and 164), which are predicted to form a disulfide bond; two conserved tryptophan residues in TM4 and TM6 (positions 140 and 316); and conserved proline residues in TM5, TM6, and TM7 (positions 186, 318, and 355, respectively). The putative TM domains are shown in Figure 1. H_4 shares sequence identity with human H_3 receptors. Sequence comparison indicates an overall sequence identity of about 44% and overall sequence similarity of about 51%.

The present invention also contemplates an assay method and system for identifying selective H_4 receptor ligands. The method involves detecting binding of a test compound to isolated cell membranes containing the histamine H_4 receptor. The assay system comprises transformed host cells that express H_4 receptors, where the number of cells in the assay system is sufficient to detect an alteration in second messenger accumulation. The test system also includes an appropriate cell culture medium to permit cell growth and viability, and preferably tissue culture plates or arrays containing the host cells in cell culture medium. In a specific embodiment, the second messenger that is detected is cAMP. In a further embodiment, the receptor is a human receptor.

The invention also discloses a method for identifying a test compound that antagonizes or agonizes histamine H_4 receptors. The method comprises detecting an increase (antagonist) or decrease (agonist) in the level of a second messenger in the assay system when contacted with the test compound.

Thus, the present invention advantageously provides H_4 protein, including fragments, derivatives, and analogs of H_4; H_4 nucleic acids, including oligonucleotide primers and probes, and H_4 regulatory sequences (especially an H_4 primer and splice sites with introns); H_4-specific antibodies; and related methods of using these materials to detect the presence of H_4

proteins or nucleic acids, H_4 binding partners, and in screens for agonists and antagonists of H_4.

The following sections of the application, which are delineated by headings (in bold) and sub-headings (in bold italics), which cover these three aspects of the invention, are provided for clarity, and not by way of limitation.

General Definitions

As used herein, the term "isolated" means that the referenced material is removed from the environment in which it is normally found. Thus, an isolated biological material can be free of cellular components, *i.e.*, components of the cells in which the material is found or produced in nature. In the case of nucleic acid molecules, an isolated nucleic acid includes a PCR product, an isolated mRNA, a cDNA, or a restriction fragment. In another embodiment, an isolated nucleic acid is preferably excised from the chromosome in which it may be found, and more preferably is no longer joined to non-regulatory, non-coding regions, or to other genes, located upstream or downstream of the gene contained by the isolated nucleic acid molecule when found in the chromosome. In yet another embodiment, the isolated nucleic acid lacks one or more introns. Isolated nucleic acid molecules include sequences inserted into plasmids, cosmids, artificial chromosomes, and the like. Thus, in a specific embodiment, a recombinant nucleic acid is an isolated nucleic acid. An isolated protein may be associated with other proteins or nucleic acids, or both, with which it associates in the cell, or with cellular membranes if it is a membrane-associated protein. A protein expressed from a vector in a cell, particularly a cell in which the protein is normally not expressed is also a regarded as isolated. An isolated organelle, cell, or tissue is removed from the anatomical site in which it is found in a cell or an organism. An isolated material may be, but need not be, purified.

The term "purified" as used herein refers to material that has been isolated under conditions that reduce or eliminate the presence of unrelated materials, *i.e.*, contaminants, including native materials from which the material is obtained. For example, a purified protein is preferably substantially free of other proteins or nucleic acids with which it is associated in a cell; a purified nucleic acid molecule is preferably substantially free of proteins or other unrelated nucleic acid molecules with which it can be found within a cell. As used herein, the term "substantially free" is used operationally, in the context of analytical testing of the material. Preferably, purified material substantially free of contaminants is at least 50% pure; more

preferably, at least 90% pure; and more preferably still at least 99% pure. Purity can be evaluated by chromatography, gel electrophoresis, immunoassay, composition analysis, biological assay, and other methods known in the art.

Methods for purification are well-known in the art. For example, nucleic acids can be purified by precipitation, chromatography (including preparative solid phase chromatography, oligonucleotide hybridization, and triple helix chromatography), ultracentrifugation, and other means. Polypeptides and proteins can be purified by various methods including, without limitation, preparative disc-gel electrophoresis, isoelectric focusing, HPLC, reversed-phase HPLC, gel filtration, ion exchange and partition chromatography, precipitation and salting-out chromatography, extraction, and countercurrent distribution. For some purposes, it is preferable to produce the polypeptide in a recombinant system in which the protein contains an additional sequence tag that facilitates purification, such as, but not limited to, a polyhistidine sequence, or a sequence that specifically binds to an antibody, such as FLAG and GST. The polypeptide can then be purified from a crude lysate of the host cell by chromatography on an appropriate solid-phase matrix. Alternatively, antibodies produced against the protein or against peptides derived therefrom can be used as purification reagents. Cells can be purified by various techniques, including centrifugation, matrix separation (*e.g.*, nylon wool separation), panning and other immunoselection techniques, depletion (*e.g.*, complement depletion of contaminating cells), and cell sorting (*e.g.*, fluorescence activated cell sorting [FACS]). Other purification methods are possible. A purified material may contain less than about 50%, preferably less than about 75%, and most preferably less than about 90%, of the cellular components with which it was originally associated. The "substantially pure" indicates the highest degree of purity which can be achieved using conventional purification techniques known in the art.

In a specific embodiment, the term "about" or "approximately" means within a scientifically acceptable error range for a given value relative to the precision with which the value is or can be measured, *e.g.*, within 20%, preferably within 10%, and more preferably within 5% of a given value or range. Alternatively, particularly with biological systems, the term can mean within an order of magnitude, preferably within 5-fold and more preferably within 2-fold of a given value.

A "sample" as used herein refers to a biological material which can be tested for

the presence of H_4 protein or H_4 nucleic acids. Such samples can be obtained from animal subjects, such as humans and non-human animals, and include tissue, especially muscle, biopsies, blood and blood products; plural effusions; cerebrospinal fluid (CSF); ascites fluid; and cell culture.

Non-human animals include, without limitation, laboratory animals such as mice, rats, rabbits, hamsters, guinea pigs, etc.; domestic animals such as dogs and cats; and, farm animals such as sheep, goats, pigs, horses, and cows.

The use of italics indicates a nucleic acid molecule; normal text indicates the polypeptide or protein.

The term "ligand" refers to a compound that recognizes and binds to a receptor binding site. In a specific embodiment, the ligand binds to the histamine H_4 receptors of the invention. Upon binding to the receptor, the ligand may produce agonist or antagonist functional effects.

The term "agonist" refers to a ligand that binds to the receptor and produces a functional effect similar to that produced by the endogenous ligand for the receptor. In a specific embodiment, the agonist at the histamine H_4 receptor produces an effect similar to that produced by histamine, the endogenous ligand (histamine) for the H_4 receptor. Examples of such agonists include, but are not limited to, R-α-methyl histamine and imetit.

The term "antagonist" refers to a ligand that binds to the receptor and blocks a functional effect produced by an agonist for the receptor or the endogenous ligand of the receptor. Examples of such antagonists include, but are not limited to, thioperamide.

The term "selective" refers to the ability of a histamine H_4 agonist or antagonist to elicit a response from the H_4 receptor while eliciting minimal responses from another receptor. Stated differently, a selective H_4 agonist may be a potent agonist for the H_4 receptor while agonizing another receptor, such as another G-protein coupled receptor and particularly another histamine receptor, poorly or not at all.

The term "ability to elicit a response" refers to the ability of a H_4 agonist or antagonist ligand to agonize or antagonize H_4 receptor activity.

As used herein the term "transformed cell" refers to a modified host cell that expresses a functional H_4 receptor expressed from a vector encoding the histamine receptor. Any cell can be used, preferably a mammalian cell, and more preferably a HEK cell.

A "functional histamine receptor" is a receptor that binds histamine or H_4 agonists and transduces a signal upon such binding. Preferably the H_4 receptor is a human H_4. Preferably, the signal that is transduced is accumulation of a second messenger, preferably cAMP. Histamine H_4 receptors may be derived from a variety of sources, including mammal, *e.g.*, human, bovine, porcine, and canine; and avian.

The cells of the invention are particularly suitable for an assay system for histamine H_4 receptor ligands that modulate second messenger accumulation. An "assay system" is one or more collections of such cells, *e.g.*, in a microwell plate or some other culture system. To permit evaluation of the effects of a test compound on the cells, the number of cells in a single assay system is sufficient to express a detectable amount of the regulated second messenger accumulation at least under conditions of maximum second messenger accumulation.

A "second messenger" is an intracellular molecule or ion, where formation and/or accumulation of the second messenger is regulated by activation of cellular membranes. In one embodiment, cellular membranes contain G-protein coupled receptor, ion channels, and tyroinse kinase receptors. In the context of this invention, the cellular membrane is a G-protein coupled receptor, preferably a histamine H_4 receptor. In a specific embodiment, the second messenger is one or more of cAMP, cGMP, inositol phosphate, DAG, and ions such as calcium and potassium. Preferably, the second messenger is cAMP.

A "test compound" or "candidate compound" is any molecule that can be tested for its ability to bind H_4 receptors, and preferably modulate second messenger accumulation through the H_4 receptor, as set forth herein. A compound that binds, and preferably modulates H_4 is a "lead compound" suitable for further testing and development as an H_4 agonist or antagonist.

As used herein, the term "provide" refers to supplying the compounds or pharmaceutical compositions of the present invention to cells or to an animal, preferably a human, in any form. For example, a prodrug form of the compounds may be provided the subject, which then is metabolized to the compound in the body.

H_4 Receptor

H_4 receptor, as defined herein, refers to a polypeptide having about 390 amino acids.

The protein is transcribed from a nucleic acid sequence that is about 1173 base pairs in length. The H_4 protein has significant homology to the H_3 receptor. Thus, H_4 refers to orthologs and allelic variants, *e.g.*, a protein having greater than about 50%, preferably greater than 80%, more preferably still greater than 90%, and even more preferably greater than 95% overall sequence identity to SEQ ID NO: 2. Allelic variants may differ from 1 to about 5 amino residues from SEQ ID NO:2. In a specific embodiment, H_4 has an amino acid sequence as shown in SEQ ID NO: 2.

Sequence comparison studies between human H_4 protein and human H_3 protein, indicates sequence identity of about 44%. The predicted protein sequence contains residues which are characteristic of the class of biogenic amine receptors. As in other biogenic amine receptors, the present protein contains conserved aspartate residues within TM2 and TM3 (at positions 61 and 94). A DY motif, found only in histamine and muscarinic receptors, is found at positions 94-95. A DRY sequence motif that exists at the cytoplasmic interface of TM3 of all GPCRs also can be found in amino acids 111-113. Also present are two conserved cysteine residues in the first and second extracellular loops (position 88 and 164), two conserved tryptophan residues in TM4 and TM6 (positions 140 and 316); and conserved proline residues in TM5, TM6, and TM7 (positions 186, 318, and 355, respectively).

H_4 receptors, like H_2 and H_3 receptors, modulates adenylyl cyclase activity. Therefore, the receptor modulates accumulation of the intracellular messenger cAMP. Modulation of H_4 receptors may be a treatment for transplanted organ rejection, asthma, allergies, and autoimmune pathologies, such as, but not limited to, multiple sclerosis, type I diabetes, and rheumatoid arthritis. Additionally, CNS functions such as cognitive and memory defects may be treated.

H_4 fragments, derivatives, and analogs can be characterized by one or more of the characteristics of H_4 protein. In a specific embodiment, in order to develop the specific C-terminal and N-terminal H_4 antibodies, antibodies can be raised against extracellular or cytoplasmic portions of the H_4 protein, preferably or antigenic peptides identified using a hydrophobicity profile or other algorithms.

Analogs and derivatives of the H_4 receptor of the invention have the same or homologous characteristics of H_4 as set forth above. For example, a truncated form of H_4 can be provided. Such a truncated form includes H_4 with a either an N-terminal, C-terminal, or internal deletion. In a specific embodiment, the derivative is functionally active, *i.e.,* capable of exhibiting one or more functional activities associated with a full-length, wild-type H_4 of the invention. Such functions include, but are not limited to, inhibition of adenylyl cyclase activity and cAMP formation. Alternatively, a H_4 chimeric fusion protein can be prepared in which the H_4 portion of the fusion protein has one or more characteristics of H_4. Such fusion proteins include fusions of the H_4 receptor with a marker polypeptide, such as FLAG, a histidine tag, a myc tag, or glutathione-S-transferase (GST). Alternatively, the H_4 receptor can be fused with an expression-related peptide, such as yeast α-mating factor, a heterogeneous signal peptide, or a peptide that renders the protein more stable upon expression. H_4 can also be fused with a unique phosphorylation site for labeling.

H_4 analogs can be made by altering encoding nucleic acid sequences by substitutions, additions or deletions that provide for functionally similar molecules, *i.e.,* molecules that perform one or more H_4 functions. In a specific embodiment, an analog of H_4 is a sequence-conservative variant of H_4. In another embodiment, an analog of H_4 is a function-conservative variant. In yet another embodiment, an analog of H_4 is an allelic variant or a homologous variant from another species. In a specific embodiment, human variants of H_4 are described.

H_4 derivatives include, but are by no means limited to, phosphorylated H_4, glycosylated H_4, methylated H_4, and other H_4 proteins that are otherwise chemically modified. H_4 derivatives also include labeled variants, *e.g.,* radio-labeled with iodine (or, as pointed out above, phosphorous); a detectable molecule, such as but by no means limited to biotin, a chelating group complexed with a metal ion, a chromophore or fluorophore, a gold colloid, or a particle such as a latex bead; or attached to a water soluble polymer.

Cloning and Expression of H_4

The present invention contemplates analysis and isolation of a gene encoding a functional or mutant H_4, including a full length, or naturally occurring form of H_4, and any antigenic fragments thereof from any source, preferably human. It further contemplates

expression of functional or mutant H_4 protein for evaluation, diagnosis, or therapy.

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, *e.g.*, Sambrook, Fritsch & Maniatis, Molecular Cloning: A Laboratory Manual, Second Edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (herein "Sambrook *et al.*, 1989"); DNA Cloning: A Practical Approach, Volumes I and II (D.N. Glover ed. 1985); Oligonucleotide Synthesis (M.J. Gait ed. 1984); Nucleic Acid Hybridization [B.D. Hames & S.J. Higgins eds. (1985)]; Transcription And Translation [B.D. Hames & S.J. Higgins, eds. (1984)]; Animal Cell Culture [R.I. Freshney, ed. (1986)]; Immobilized Cells And Enzymes [IRL Press, (1986)]; B.Perbal, A Practical Guide To Molecular Cloning (1984); F.M. Ausubel *et al.* (eds.), Current Protocols in Molecular Biology, John Wiley & Sons, Inc. (1994).

Molecular Biology - Definitions

"Amplification" of DNA as used herein denotes the use of polymerase chain reaction (PCR) to increase the concentration of a particular DNA sequence within a mixture of DNA sequences. For a description of PCR see Saiki *et al.*, Science, 239:487, 1988.

As used herein, "sequence-specific oligonucleotides" refers to related sets of oligonucleotides that can be used to detect allelic variations or mutations in the H_4 gene.

The nucleic acid molecules (polynucleotides) herein may be flanked by natural regulatory (expression control) sequences, or may be associated with heterologous sequences, including promoters, internal ribosome entry sites (IRES) and other ribosome binding site sequences, enhancers, response elements, suppressors, signal sequences, polyadenylation sequences, introns, 5'- and 3'- non-coding regions, and the like. The nucleic acids may also be modified by many means known in the art. Non-limiting examples of such modifications include methylation, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, and internucleotide modifications such as, for example, those with uncharged linkages (*e.g.*, methyl phosphonates, phosphotriesters, phosphoroamidates, carbamates, etc.) and with charged linkages (*e.g.*, phosphorothioates, phosphorodithioates, etc.). Polynucleotides may contain one or more additional covalently linked moieties, such as, for example, proteins (*e.g.*, nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.), intercalators (*e.g.*, acridine, psoralen, etc.), chelators (*e.g.*, metals, radioactive metals, iron, oxidative metals, etc.), and

alkylators. The polynucleotides may be derivatized by formation of a methyl or ethyl phosphotriester or an alkyl phosphoramidate linkage. Furthermore, the polynucleotides herein may also be modified with a label capable of providing a detectable signal, either directly or indirectly. Exemplary labels include radioisotopes, fluorescent molecules, biotin, and the like.

A "coding sequence" or a sequence "encoding" an expression product, such as a RNA, polypeptide, protein, or enzyme, is a nucleotide sequence that, when expressed, results in the production of that RNA, polypeptide, protein, or enzyme, *i.e.*, the nucleotide sequence encodes an amino acid sequence for that polypeptide, protein or enzyme. A coding sequence for a protein may include a start codon (usually ATG) and a stop codon.

The term "gene", also called a "structural gene" means a DNA sequence that codes for or corresponds to a particular sequence of amino acids which comprise all or part of one or more proteins or enzymes, and may or may not include introns and regulatory DNA sequences, such as promoter sequences, 5'-untranslated region, or 3'-untranslated region which affect for example the conditions under which the gene is expressed.

A "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site (conveniently defined for example, by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase. The present invention includes the H_4 receptor gene promoter found in the genome, which can be operatively associated with a H_4 coding sequence with a heterologous coding sequence.

The term "host cell" means any cell of any organism that is selected, modified, transformed, grown, or used or manipulated in any way, for the production of a substance by the cell, for example the expression by the cell of a gene, a DNA or RNA sequence, a protein or an enzyme. Host cells can further be used for screening or other assays, as described *infra*.

A coding sequence is "under the control of" or "operatively associated with" transcriptional and translational control sequences in a cell when RNA polymerase transcribes the coding sequence into mRNA, which is then trans-RNA spliced (if it contains introns) and

translated, in the case of mRNA, into the protein encoded by the coding sequence.

The terms "express" and "expression" mean allowing or causing the information in a gene or DNA sequence to become manifest, for example producing a protein by activating the cellular functions involved in transcription and translation of a corresponding gene or DNA sequence. A DNA sequence is expressed in or by a cell to form an "expression product" such as a protein. The expression product itself , *e.g.*, the resulting protein, may also be said to be "expressed" by the cell. An expression product can be characterized as intracellular, extracellular, transmembrane, or secreted. The term "intracellular" means something that is inside a cell. The term "extracellular" means something that is outside a cell. The H_4 receptor is a seven transmembrane protein with intracellular and extracellular domains. A substance is "secreted" by a cell if it appears in significant measure outside the cell, from somewhere on or inside the cell.

The term "transfection" means the introduction of a foreign nucleic acid into a cell. The term "transformation" means the introduction of a "foreign" (*i.e.*, extrinsic or extracellular) gene, DNA or RNA sequence to a host cell, so that the host cell will express the introduced gene or sequence to produce a desired substance, typically a protein or enzyme coded by the introduced gene or sequence. The introduced gene or sequence may also be called a "cloned" or "foreign" gene or sequence, may include regulatory or control sequences, such as start, stop, promoter, signal, secretion, or other sequences used by a cell's genetic machinery. The gene or sequence may include nonfunctional sequences or sequences with no known function. A host cell that receives and expresses introduced DNA or RNA has been "transformed" and is a "transformant" or a "clone." The DNA or RNA introduced to a host cell can come from any source, including cells of the same genus or species as the host cell, or cells of a different genus or species.

The terms "vector", "cloning vector" and "expression vector" mean the vehicle by which a DNA or RNA sequence (*e.g.*, a foreign gene) can be introduced into a host cell, so as to transform the host and promote expression (*e.g.*, transcription and translation) of the introduced sequence. Vectors include plasmids, phages, viruses, etc.; they are discussed in greater detail below.

Vectors typically comprise the DNA of a transmissible agent, into which foreign DNA is inserted. A common way to insert one segment of DNA into another segment of DNA

involves the use of enzymes called restriction enzymes that cleave DNA at specific sites (specific groups of nucleotides) called restriction sites. A "cassette" refers to a DNA coding sequence or segment of DNA that codes for an expression product that can be inserted into a vector at defined restriction sites. The cassette restriction sites are designed to ensure insertion of the cassette in the proper reading frame. Generally, foreign DNA is inserted at one or more restriction sites of the vector DNA, and then is carried by the vector into a host cell along with the transmissible vector DNA. A segment or sequence of DNA having inserted or added DNA, such as an expression vector, can also be called a "DNA construct." A common type of vector is a "plasmid", which generally is a self-contained molecule of double-stranded DNA, usually of bacterial origin, that can readily accept additional (foreign) DNA and which can readily introduced into a suitable host cell. A plasmid vector often contains coding DNA and promoter DNA and has one or more restriction sites suitable for inserting foreign DNA. Coding DNA is a DNA sequence that encodes a particular amino acid sequence for a particular protein or enzyme. Promoter DNA is a DNA sequence which initiates, regulates, or otherwise mediates or controls the expression of the coding DNA. Promoter DNA and coding DNA may be from the same gene or from different genes, and may be from the same or different organisms. A large number of vectors, including plasmid and fungal vectors, have been described for replication and/or expression in a variety of eukaryotic and prokaryotic hosts. Non-limiting examples include pKK plasmids (Clonetech), pUC plasmids, pET plasmids (Novagen, Inc., Madison, WI), pRSET or pREP plasmids (Invitrogen, San Diego, CA), or pMAL plasmids (New England Biolabs, Beverly, MA), and many appropriate host cells, using methods disclosed or cited herein or otherwise known to those skilled in the relevant art. Recombinant cloning vectors will often include one or more replication systems for cloning or expression, one or more markers for selection in the host, *e.g.*, antibiotic resistance, and one or more expression cassettes.

The term "expression system" means a host cell and compatible vector under suitable conditions, *e.g.*, for the expression of a protein coded for by foreign DNA carried by the vector and introduced to the host cell. Common expression systems include E. coli host cells and plasmid vectors, insect host cells and Baculovirus vectors, and mammalian host cells and vectors.

The term "heterologous" refers to a combination of elements not naturally occurring. For example, heterologous DNA refers to DNA not naturally located in the cell, or

in a chromosomal site of the cell. Preferably, the heterologous DNA includes a gene foreign to the cell. A heterologous expression regulatory element is such an element operatively associated with a different gene than the one it is operatively associated with in nature. In the context of the present invention, an H_4 gene is heterologous to the vector DNA in which it is inserted for cloning or expression, and it is heterologous to a host cell containing such a vector, in which it is expressed, *e.g.*, a HEK cell.

The terms "mutant" and "mutation" mean any detectable change in genetic material, *e.g.*, DNA, or any process, mechanism, or result of such a change. This includes gene mutations, in which the structure (*e.g.*, DNA sequence) of a gene is altered, any gene or DNA arising from any mutation process, and any expression product (*e.g.*, protein or enzyme) expressed by a modified gene or DNA sequence. The term "variant" may also be used to indicate a modified or altered gene, DNA sequence, enzyme, cell, etc., *i.e.*, any kind of mutant.

"Sequence-conservative variants" of a polynucleotide sequence are those in which a change of one or more nucleotides in a given codon position results in no alteration in the amino acid encoded at that position.

"Function-conservative variants" are those in which a given amino acid residue in a protein or enzyme has been changed without altering the overall conformation and function of the polypeptide, including, but not limited to, replacement of an amino acid with one having similar properties (such as, for example, polarity, hydrogen bonding potential, acidic, basic, hydrophobic, aromatic, and the like). Amino acids with similar properties are well known in the art. For example, arginine, histidine and lysine are hydrophilic-basic amino acids and may be interchangeable. Similarly, isoleucine, a hydrophobic amino acid, may be replaced with leucine, methionine or valine. Such changes are expected to have little or no effect on the apparent molecular weight or isoelectric point of the protein or polypeptide. Amino acids other than those indicated as conserved may differ in a protein or enzyme so that the percent protein or amino acid sequence similarity between any two proteins of similar function may vary and may be, for example, from 70% to 99% as determined according to an alignment scheme such as by the Cluster Method, wherein similarity is based on the MEGALIGN algorithm. A "function-conservative variant" also includes a polypeptide or enzyme which has at least 60 % amino acid identity as determined by BLAST or FASTA algorithms, preferably at least 75%, most preferably at least 85%, and even more preferably at least 90%, and which has the same or

substantially similar properties or functions as the native or parent protein or enzyme to which it is compared.

As used herein, the term "homologous" in all its grammatical forms and spelling variations refers to the relationship between proteins that possess a "common evolutionary origin," including proteins from superfamilies (*e.g.*, the immunoglobulin superfamily) and homologous proteins from different species (*e.g.*, myosin light chain, etc.) (Reeck *et al.*, Cell 1987, 50:667). Such proteins (and their encoding genes) have sequence homology, as reflected by their sequence similarity, whether in terms of percent similarity or the presence of specific residues or motifs at conserved positions.

Accordingly, the term "sequence similarity" in all its grammatical forms refers to the degree of identity or correspondence between nucleic acid or amino acid sequences of proteins that may or may not share a common evolutionary origin (see Reeck *et al.*, *supra*). However, in common usage and in the instant application, the term "homologous," when modified with an adverb such as "highly," may refer to sequence similarity and may or may not relate to a common evolutionary origin.

In a specific embodiment, two DNA sequences are "substantially homologous" or "substantially similar" when at least about 80%, and most preferably at least about 90 or 95% of the nucleotides match over the defined length of the DNA sequences, as determined by sequence comparison algorithms, such as BLAST, FASTA, DNA Strider, etc. An example of such a sequence is an allelic or species variant of the specific H_4 gene of the invention. Sequences that are substantially homologous can be identified by comparing the sequences using standard software available in sequence data banks, or in a Southern hybridization experiment under, for example, stringent conditions as defined for that particular system.

Similarly, in a particular embodiment, two amino acid sequences are "substantially homologous" or "substantially similar" when greater than 80% of the amino acids are identical, or greater than about 90% are similar (functionally identical). Preferably, the similar or homologous sequences are identified by alignment using, for example, the GCG (Genetics Computer Group, Program Manual for the GCG Package, Version 7, Madison, Wisconsin) pileup program, or any of the programs described above (BLAST, FASTA, etc)

A nucleic acid molecule is "hybridizable" to another nucleic acid molecule, such as a cDNA, genomic DNA, or RNA, when a single stranded form of the nucleic acid molecule

can anneal to the other nucleic acid molecule under the appropriate conditions of temperature and solution ionic strength (see Sambrook *et al.*, *supra*). The conditions of temperature and ionic strength determine the "stringency" of the hybridization. For preliminary screening for homologous nucleic acids, low stringency hybridization conditions, corresponding to a Tm (melting temperature) of 55°C, can be used, *e.g.*, 5x SSC, 0.1% SDS, 0.25% milk, and no formamide; or 30% formamide, 5x SSC, 0.5% SDS. Moderate stringency hybridization conditions correspond to a higher Tm, *e.g.*, 40% formamide, with 5x or 6x SSC. High stringency hybridization conditions correspond to the highest Tm, *e.g.*, 50% formamide, 5x or 6x SSC. SSC is a 0.15M NaCl, 0.015M Na-citrate buffer. Hybridization requires that the two nucleic acids contain complementary sequences, although depending on the stringency of the hybridization, mismatches between bases are possible. The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementation, variables well known in the art. The greater the degree of similarity or homology between two nucleotide sequences, the greater the value of Tm for hybrids of nucleic acids having those sequences. The relative stability (corresponding to higher Tm) of nucleic acid hybridizations decreases in the following order: RNA:RNA, DNA:RNA, DNA:DNA. For hybrids of greater than 100 nucleotides in length, equations for calculating Tm have been derived (see Sambrook *et al.*, *supra*, 9.50-9.51). For hybridization with shorter nucleic acids, *i.e.*, oligonucleotides, the position of mismatches becomes more important, and the length of the oligonucleotide determines its specificity (see Sambrook *et al.*, *supra*, 11.7-11.8). A minimum length for a hybridizable nucleic acid is at least about 10 nucleotides; preferably at least about 15 nucleotides; and more preferably the length is at least about 20 nucleotides.

In a specific embodiment, the term "standard hybridization conditions" refers to a Tm of 55°C, and utilizes conditions as set forth above. In a preferred embodiment, the Tm is 60°C; in a more preferred embodiment, the Tm is 65°C. In a specific embodiment, "high stringency" refers to hybridization and/or washing conditions at 68°C in 0.2xSSC, at 42°C in 50% formamide, 4xSSC, or under conditions that afford levels of hybridization equivalent to those observed under either of these two conditions.

As used herein, the term "oligonucleotide" refers to a nucleic acid, generally of at least 10, preferably at least 15, and more preferably at least 20 nucleotides, preferably no more than 100 nucleotides, that is hybridizable to a genomic DNA molecule, a cDNA molecule, or an

mRNA molecule encoding a gene, mRNA, cDNA, or other nucleic acid of interest. Oligonucleotides can be labeled, *e.g.*, with ^{32}P-nucleotides or nucleotides to which a label, such as biotin, has been covalently conjugated. In one embodiment, a labeled oligonucleotide can be used as a probe to detect the presence of a nucleic acid. In another embodiment, oligonucleotides (one or both of which may be labeled) can be used as PCR primers, either for cloning full length or a fragment of H_4, or to detect the presence of nucleic acids encoding H_4. In a further embodiment, an oligonucleotide of the invention can form a triple helix with a H_4 DNA molecule. Generally, oligonucleotides are prepared synthetically, preferably on a nucleic acid synthesizer. Accordingly, oligonucleotides can be prepared with non-naturally occurring phosphoester analog bonds, such as thioester bonds, etc.

The present invention provides antisense nucleic acids (including ribozymes), which may be used to inhibit expression of H_4 of the invention. Inhibition of H_4 expression may be desired when upregulation of H_4 receptor expression or excessive inhibition of cAMP formation induces disease states such as, transplant organ rejection; asthma; allergies; autoimmune pathologies, such as, but not limited to, multiple sclerosis, type I diabetes, and rheumatoid arthritis; and CNS functions such as cognitive and memory defects. An "antisense nucleic acid" is a single stranded nucleic acid molecule which, on hybridizing under cytoplasmic conditions with complementary bases in an RNA or DNA molecule, inhibits the latter's role. If the RNA is a messenger RNA transcript, the antisense nucleic acid is a countertranscript or mRNA-interfering complementary nucleic acid. As presently used, "antisense" broadly includes RNA-RNA interactions, RNA-DNA interactions, ribozymes and RNase-H mediated arrest. Antisense nucleic acid molecules can be encoded by a recombinant gene for expression in a cell (*e.g.*, U.S. Patent No. 5,814,500; U.S. Patent No. 5,811,234), or alternatively they can be prepared synthetically (*e.g.*, U.S. Patent No. 5,780,607).

Specific non-limiting examples of synthetic oligonucleotides envisioned for this invention include oligonucleotides that contain phosphorothioates, phosphotriesters, methyl phosphonates, short chain alkyl, or cycloalkyl intersugar linkages or short chain heteroatomic or heterocyclic intersugar linkages. Most preferred are those with CH_2-NH-O-CH_2, CH_2-N(CH_3)-O-CH_2, CH_2-O-N(CH_3)-CH_2, CH_2-N(CH_3)-N(CH_3)-CH_2 and O-N(CH_3)-CH_2-CH_2 backbones (where phosphodiester is O-PO_2-O-CH_2). U.S. Patent No. 5,677,437 describes heteroaromatic oligonucleoside linkages. Nitrogen linkers or groups containing nitrogen can also be used to

prepare oligonucleotide mimics (U.S. Patents No. 5,792,844 and No. 5,783,682). U.S. Patent No. 5,637,684 describes phosphoramidate and phosphorothioamidate oligomeric compounds. Also envisioned are oligonucleotides having morpholino backbone structures (U.S. Patent No. 5,034,506). In other embodiments, such as the peptide-nucleic acid (PNA) backbone, the phosphodiester backbone of the oligonucleotide may be replaced with a polyamide backbone, the bases being bound directly or indirectly to the aza nitrogen atoms of the polyamide backbone (Nielsen *et al.*, Science 254:1497, 1991). Other synthetic oligonucleotides may contain substituted sugar moieties comprising one of the following at the 2' position: OH, SH, SCH_3, F, OCN, $O(CH_2)_nNH_2$ or $O(CH_2)_nCH_3$ where n is from 1 to about 10; C1 to C10 lower alkyl, substituted lower alkyl, alkaryl or aralkyl; Cl; Br; CN; CF_3; OCF_3; O-; S-, or N-alkyl; O-, S-, or N-alkenyl; $SOCH_3$; SO_2CH_3; ONO_2; NO_2; N_3; NH_2; heterocycloalkyl; heterocycloalkaryl; aminoalkylamino; polyalkylamino; substitued silyl; a fluorescein moiety; an RNA cleaving group; a reporter group; an intercalator; a group for improving the pharmacokinetic properties of an oligonucleotide; or a group for improving the pharmacodynamic properties of an oligonucleotide, and other substituents having similar properties. Oligonucleotides may also have sugar mimetics such as cyclobutyls or other carbocyclics in place of the pentofuranosyl group. Nucleotide units having nucleosides other than adenosine, cytidine, guanosine, thymidine and uridine, such as inosine, may be used in an oligonucleotide molecule.

H_4 Nucleic Acids

A gene encoding H_4, whether genomic DNA or cDNA, can be isolated from any source, particularly from a human cDNA or genomic library. Methods for obtaining H_4 gene are well known in the art, as described above (see, *e.g.*, Sambrook *et al.*, 1989, *supra*). The DNA may be obtained by standard procedures known in the art from cloned DNA (*e.g.*, a DNA "library"), and preferably is obtained from a cDNA library prepared from tissues with high level expression of the protein, by chemical synthesis, by cDNA cloning, or by the cloning of genomic DNA, or fragments thereof, purified from the desired cell (See, for example, Sambrook *et al.*, 1989, *supra*; Glover, D.M. (ed.), 1985, DNA Cloning: A Practical Approach, MRL Press, Ltd., Oxford, U.K. Vol. I, II). Clones derived from genomic DNA may contain regulatory and intron DNA regions in addition to coding regions; clones derived from cDNA will not contain intron sequences. Whatever the source, the gene should be molecularly cloned into a suitable vector for propagation of the gene. Identification of the specific DNA fragment containing the desired

H_4 gene may be accomplished in a number of ways. For example, a portion of a H_4 gene exemplified *infra* can be purified and labeled to prepare a labeled probe, and the generated DNA may be screened by nucleic acid hybridization to the labeled probe (Benton and Davis, Science 1977, 196:180; Grunstein and Hogness, Proc. Natl. Acad. Sci. U.S.A. 1975, 72:3961). Those DNA fragments with substantial homology to the probe, such as an allelic variant from another individual, will hybridize. In a specific embodiment, highest stringency hybridization conditions are used to identify a homologous H_4 gene.

Further selection can be carried out on the basis of the properties of the gene, *e.g.*, if the gene encodes a protein product having the isoelectric, electrophoretic, amino acid composition, partial or complete amino acid sequence, antibody binding activity, or ligand binding profile of H_4 protein as disclosed herein. Thus, the presence of the gene may be detected by assays based on the physical, chemical, immunological, or functional properties of its expressed product.

Other DNA sequences which encode substantially the same amino acid sequence as a H_4 gene may be used in the practice of the present invention. These include but are not limited to allelic variants, species variants, sequence conservative variants, and functional variants.

Amino acid substitutions may also be introduced to substitute an amino acid with a particularly preferable property. For example, a Cys may be introduced a potential site for disulfide bridges with another Cys.

The genes encoding H_4 derivatives and analogs of the invention can be produced by various methods known in the art. The manipulations which result in their production can occur at the gene or protein level. For example, the cloned H_4 gene sequence can be modified by any of numerous strategies known in the art (Sambrook *et al.*, 1989, *supra*). The sequence can be cleaved at appropriate sites with restriction endonuclease(s), followed by further enzymatic modification if desired, isolated, and ligated *in vitro*. In the production of the gene encoding a derivative or analog of H_4, care should be taken to ensure that the modified gene remains within the same translational reading frame as the H_4 gene, uninterrupted by translational stop signals, in the gene region where the desired activity is encoded.

Additionally, the v-encoding nucleic acid sequence can be mutated *in vitro* or *in vivo*, to create and/or destroy translation, initiation, and/or termination sequences, or to create

variations in coding regions and/or form new restriction endonuclease sites or destroy preexisting ones, to facilitate further *in vitro* modification. Such modifications can be made to introduce restriction sites and facilitate cloning the H_4 gene into an expression vector. Any technique for mutagenesis known in the art can be used, including but not limited to, *in vitro* site-directed mutagenesis (Hutchinson, C., *et al.*, J. Biol. Chem. 253:6551, 1978; Zoller and Smith, DNA 3:479-488, 1984; Oliphant *et al.*, Gene 44:177, 1986; Hutchinson *et al.*, Proc. Natl. Acad. Sci. U.S.A. 83:710, 1986), use of TAB™ linkers (Pharmacia), etc. PCR techniques are preferred for site directed mutagenesis (see Higuchi, 1989, "Using PCR to Engineer DNA", in PCR Technology: Principles and Applications for DNA Amplification, H. Erlich, ed., Stockton Press, Chapter 6, pp. 61-70).

The identified and isolated gene can then be inserted into an appropriate cloning vector. A large number of vector-host systems known in the art may be used. Possible vectors include, but are not limited to, plasmids or modified viruses, but the vector system must be compatible with the host cell used. Examples of vectors include, but are not limited to, E. coli, bacteriophages such as lambda derivatives, or plasmids such as Bluescript, pBR322 derivatives or pUC plasmid derivatives, *e.g.*, pGEX vectors, pmal-c, pFLAG, etc. The insertion into a cloning vector can, for example, be accomplished by ligating the DNA fragment into a cloning vector which has complementary cohesive termini. However, if the complementary restriction sites used to fragment the DNA are not present in the cloning vector, the ends of the DNA molecules may be enzymatically modified. Alternatively, any site desired may be produced by ligating nucleotide sequences (linkers) onto the DNA termini; these ligated linkers may comprise specific chemically synthesized oligonucleotides encoding restriction endonuclease recognition sequences. In addition, simple PCR or overlapping PCR may be used to insert a fragment into a cloning vector.

Recombinant molecules can be introduced into host cells via transformation, transfection, infection, electroporation, etc., so that many copies of the gene sequence are generated. Preferably, the cloned gene is contained on a shuttle vector plasmid, which provides for expansion in a cloning cell, *e.g.*, E. coli, and facile purification for subsequent insertion into an appropriate expression cell line, if such is desired. For example, a shuttle vector, which is a vector that can replicate in more than one type of organism, can be prepared for replication in both E. coli and Saccharomyces cerevisiae by linking sequences from an E. coli plasmid with

sequences form the yeast 2μ plasmid.

H_4 Regulatory Nucleic Acids

Elements of the H_4 promoter can be identified by scanning the human genomic region upstream of the H_4 start site, *e.g.*, by creating deletion mutants and checking for expression, or with the TRANSFAC algorithm. Sequences up to about 6 kilobases (kb) or more upstream from the H_4 start site can contain tissue-specific regulatory elements.

The term "H_4 promoter" encompasses artificial promoters. Such promoters can be prepared by deleting nonessential intervening sequences from the upstream region of the H_4 promoter, or by joining upstream regulatory elements from the H_4 promoter with a heterologous minimal promoter, such as the CMV immediate early promoter.

An H_4 promoter can be operably associated with a heterogenous coding sequence, *e.g.*, for reporter gene (luciferase and green fluorescent proteins are examples of reporter genes) in a construct. This construct will result in expression of the heterologous coding sequence under control the H_4 promoter, *e.g.*, a reporter gene can be expressed, under conditions that under normal conditions cause H_4 expression. This construct can be used in screening assays, described below, for H_4 agonists and antagonists.

Expression of H_4 Polypeptides

The nucleotide sequence coding for H_4, or antigenic fragment, derivative or analog thereof, or a functionally active derivative, including a chimeric protein, thereof, can be inserted into an appropriate expression vector, *i.e.*, a vector which contains the necessary elements for the transcription and translation of the inserted protein-coding sequence. Thus, a nucleic acid encoding H_4 of the invention can be operationally associated with a promoter in an expression vector of the invention. Both cDNA and genomic sequences can be cloned and expressed under control of such regulatory sequences. Such vectors can be used to express functional or functionally inactivated H_4 polypeptides.

The necessary transcriptional and translational signals can be provided on a recombinant expression vector, or they may be supplied by the native gene encoding H_4 and/or its flanking regions.

Potential host-vector systems include but are not limited to mammalian cell systems transfected with expression plasmids or infected with virus (*e.g.*, vaccinia virus, adenovirus, adeno-associated virus, herpes virus, etc.); insect cell systems infected with virus

(*e.g.*, baculovirus); microorganisms such as yeast containing yeast vectors; or bacteria transformed with bacteriophage, DNA, plasmid DNA, or cosmid DNA. The expression elements of vectors vary in their strengths and specificities. Depending on the host-vector system utilized, any one of a number of suitable transcription and translation elements may be used.

Expression of H_4 protein may be controlled by any promoter/enhancer element known in the art, but these regulatory elements must be functional in the host selected for expression. Promoters which may be used to control H_4 gene expression include, but are not limited to, cytomegalovirus (CMV) promoter (U.S. Patent Nos. 5,385,839 and 5,168,062), the SV40 early promoter region (Benoist and Chambon, 1981, Nature 290:304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto, *et al.*, Cell 22:787-797, 1980), the herpes thymidine kinase promoter (Wagner *et al.*, Proc. Natl. Acad. Sci. U.S.A. 78:1441-1445, 1981), the regulatory sequences of the metallothionein gene (Brinster *et al.*, Nature 296:39-42, 1982); prokaryotic expression vectors such as the β-lactamase promoter (Villa-Komaroff, *et al.*, Proc. Natl. Acad. Sci. U.S.A. 75:3727-3731, 1978), or the tac promoter (DeBoer, *et al.*, Proc. Natl. Acad. Sci. U.S.A. 80:21-25, 1983); see also "Useful proteins from recombinant bacteria" in Scientific American, 242:74-94, 1980; promoter elements from yeast or other fungi such as the Gal 4 promoter, the ADC (alcohol dehydrogenase) promoter, PGK (phosphoglycerol kinase) promoter, alkaline phosphatase promoter; and transcriptional control regions that exhibit tissue specificity, particularly endothelial cell-specific promoters.

Solubilized forms of the protein can be obtained by solubilizing inclusion bodies or reconstituting membrane components, *e.g.*, by treatment with detergent, and if desired sonication or other mechanical processes, as described above. The solubilized protein can be isolated using various techniques, such as polyacrylamide gel electrophoresis (PAGE), isoelectric focusing, 2-dimensional gel electrophoresis, chromatography (*e.g.*, ion exchange, affinity, immunoaffinity, and sizing column chromatography), centrifugation, differential solubility, immunoprecipitation, or by any other standard technique for the purification of proteins.

Vectors

A wide variety of host/expression vector combinations may be employed in expressing the DNA sequences of this invention. Useful expression vectors, for example, may consist of segments of chromosomal, non-chromosomal and synthetic DNA sequences. Suitable vectors include derivatives of SV40 and known bacterial plasmids, *e.g.*, E. coli plasmids col El,

pCR1, pBR322, pMal-C2, pET, pGEX (Smith *et al.*, Gene 67:31-40, 1988), pMB9 and their derivatives, plasmids such as RP4; phage DNAS, *e.g.*, the numerous derivatives of phage l, *e.g.*, NM989, and other phage DNA, *e.g.*, M13 and filamentous single stranded phage DNA; yeast plasmids such as the 2μ plasmid or derivatives thereof; vectors useful in eukaryotic cells, such as vectors useful in insect or mammalian cells; vectors derived from combinations of plasmids and phage DNAs, such as plasmids that have been modified to employ phage DNA or other expression control sequences; and the like.

Viral vectors, such as lentiviruses, retroviruses, herpes viruses, adenoviruses, adeno-associated viruses, vaccinia virus, baculovirus, alphavirus, and other recombinant viruses with desirable cellular tropism are also useful. Thus, a gene encoding a functional or mutant H_4 protein or polypeptide domain fragment thereof can be introduced *in vivo*, *ex vivo*, or *in vitro* using a viral vector or through direct introduction of DNA. Expression in targeted tissues can be effected by targeting the transgenic vector to specific cells, such as with a viral vector or a receptor ligand, or by using a tissue-specific promoter, or both. Targeted gene delivery is described in International Patent Publication WO 95/28494, published October 1995.

Viral vectors commonly used for *in vivo* or *ex vivo* targeting and therapy procedures are DNA-based vectors and retroviral vectors. Methods for constructing and using viral vectors are known in the art (see, *e.g.*, Miller and Rosman, BioTechniques 1992, 7:980-990). Preferably, the viral vectors are replication defective, that is, they are unable to replicate autonomously in the target cell. In general, the genome of the replication defective viral vectors which are used within the scope of the present invention lack at least one region which is necessary for the replication of the virus in the infected cell. These regions can either be eliminated (in whole or in part) or be rendered non-functional by any technique known to a person skilled in the art. These techniques include the total removal, substitution (by other sequences, in particular by the inserted nucleic acid), partial deletion or addition of one or more bases to an essential (for replication) region. Such techniques may be performed *in vitro* (on the isolated DNA) or *in situ*, using the techniques of genetic manipulation or by treatment with mutagenic agents. Preferably, the replication defective virus retains the sequences of its genome which are necessary for encapsidating the viral particles.

DNA viral vectors include an attenuated or defective DNA virus, such as but not limited to herpes simplex virus (HSV), papillomavirus, Epstein Barr virus (EBV), adenovirus,

adeno-associated virus (AAV), and the like. Defective viruses, which entirely or almost entirely lack viral genes, are preferred. Defective virus is not infective after introduction into a cell. Use of defective viral vectors allows for administration to cells in a specific, localized area, without concern that the vector can infect other cells. Thus, a specific tissue can be specifically targeted. Examples of particular vectors include, but are not limited to, a defective herpes virus 1 (HSV1) vector (Kaplitt *et al.*, Molec. Cell. Neurosci. 2:320-330, 1991), defective herpes virus vector lacking a glyco-protein L gene (Patent Publication RD 371005 A), or other defective herpes virus vectors (International Patent Publication No. WO 94/21807, published September 29, 1994; International Patent Publication No. WO 92/05263, published April 2, 1994); an attenuated adenovirus vector, such as the vector described by Stratford-Perricaudet *et al.* (J. Clin. Invest. 90:626-630, 1992; see also La Salle *et al.*, Science 259:988-990, 1993); and a defective adeno-associated virus vector (Samulski *et al.*, J. Virol. 61:3096-3101, 1987; Samulski *et al.*, J. Virol. 63:3822-3828, 1989; Lebkowski *et al.*, Mol. Cell. Biol. 8:3988-3996, 1988).

Various companies produce viral vectors commercially, including but by no means limited to Avigen, Inc. (Alameda, CA; AAV vectors), Cell Genesys (Foster City, CA; retroviral, adenoviral, AAV vectors, and lentiviral vectors), Clontech (retroviral and baculoviral vectors), Genovo, Inc. (Sharon Hill, PA; adenoviral and AAV vectors), Genvec (adenoviral vectors), IntroGene (Leiden, Netherlands; adenoviral vectors), Molecular Medicine (retroviral, adenoviral, AAV, and herpes viral vectors), Norgen (adenoviral vectors), Oxford BioMedica (Oxford, United Kingdom; lentiviral vectors), and Transgene (Strasbourg, France; adenoviral, vaccinia, retroviral, and lentiviral vectors).

Preferably, for *in vivo* administration, an appropriate immunosuppressive treatment is employed in conjunction with the viral vector, *e.g.*, adenovirus vector, to avoid immuno-deactivation of the viral vector and transfected cells. For example, immunosuppressive cytokines, such as interleukin-12 (IL-12), interferon-γ (IFN-γ), or anti-CD4 antibody, can be provided to block humoral or cellular immune responses to the viral vectors (see, *e.g.*, Wilson, Nature Medicine, 1995). In that regard, it is advantageous to employ a viral vector that is engineered to express a minimal number of antigens.

In another embodiment, the vector can be introduced *in vivo* by lipofection, as naked DNA, or with other transfection facilitating agents (peptides, polymers, etc.). Synthetic cationic lipids can be used to prepare liposomes for *in vivo* transfection of a gene encoding a

marker (Felgner, et. al., Proc. Natl. Acad. Sci. U.S.A. 84:7413-7417, 1987; Felgner and Ringold, Science 337:387-388, 1989; see Mackey, *et al.*, Proc. Natl. Acad. Sci. U.S.A. 85:8027-8031, 1988; Ulmer, *et al.*, Science 259:1745-1748, 1993). Useful lipid compounds and compositions for transfer of nucleic acids are described in International Patent Publications WO 95/18863 and WO 96/17823, and in U.S. Patent No. 5,459,127. Lipids may be chemically coupled to other molecules for the purpose of targeting (see Mackey, *et al.*, *supra*). Targeted peptides, *e.g.*, hormones or neurotransmitters, and proteins such as antibodies, or non-peptide molecules could be coupled to liposomes chemically.

Other molecules are also useful for facilitating transfection of a nucleic acid *in vivo*, such as a cationic oligopeptide (*e.g.*, International Patent Publication WO 95/21931), peptides derived from DNA binding proteins (*e.g.*, International Patent Publication WO 96/25508), or a cationic polymer (*e.g.*, International Patent Publication WO95/21931).

Alternatively, non-viral DNA vectors for gene therapy can be introduced into the desired host cells by methods known in the art, *e.g.*, electroporation, microinjection, cell fusion, DEAE dextran, calcium phosphate precipitation, use of a gene gun (ballistic transfection; see, *e.g.*, U.S. Pat. No. 5,204,253, U.S. Pat. No. 5,853,663, U.S. Pat. No. 5,885,795, and U.S. Pat. No. 5,702,384 and see Sanford, TIB-TECH, 6:299-302, 1988; Fynan *et al.*, Proc. Natl. Acad. Sci. U.S.A., 90:11478-11482, 1993; and Yang *et al.*, Proc. Natl. Acad. Sci. U.S.A., 87:1568-9572, 1990), or use of a DNA vector transporter (see, *e.g.*, Wu, *et al.*, J. Biol. Chem. 267:963-967, 1992; Wu and Wu, J. Biol. Chem. 263:14621-14624, 1988; Hartmut, *et al.*, Canadian Patent Application No. 2,012,311, filed March 15, 1990; Williams, *et al.*, Proc. Natl. Acad. Sci. USA 88:2726-2730, 1991). Receptor-mediated DNA delivery approaches can also be used (Curiel, *et al.*, Hum. Gene Ther. 3:147-154, 1992; Wu and Wu, J. Biol. Chem. 262:4429-4432, 1987). U.S. Patent Nos. 5,580,859 and 5,589,466 disclose delivery of exogenous DNA sequences, free of transfection facilitating agents, in a mammal. Recently, a relatively low voltage, high efficiency *in vivo* DNA transfer technique, termed electrotransfer, has been described (Mir, *et al.*, C.P. Acad. Sci., 321:893, 1998; WO 99/01157; WO 99/01158; WO 99/01175).

H_4 Ligands and Binding Partners

The present invention further permits identification of physiological ligands and binding partners of H_4. One method for evaluating and identifying H_4 binding partners is the

yeast two-hybrid screen. Preferably, the yeast two-hybrid screen is performed using an cell library with yeast that are transformed with recombinant H_4. Alternatively, H_4 can be used as a capture or affinity purification reagent. In another alternative, labeled H_4 can be used as a probe for binding, *e.g.*, by immunoprecipitation or Western analysis. Expected H_4 binding partners are G-proteins.

Generally, binding interactions between H_4 and any of its binding partners will be strongest under conditions approximating those found in the cytoplasm, *i.e.*, physiological conditions of ionic strength, pH and temperature. Perturbation of these conditions will tend to disrupt the stability of a binding interaction.

Antibodies to H_4

Antibodies to H_4 are useful, *inter alia*, for diagnostics and intracellular regulation of H_4 activity, as set forth below. According to the invention, a H_4 polypeptide produced recombinantly or by chemical synthesis, and fragments or other derivatives or analogs thereof, including fusion proteins, may be used as immunogens to generate antibodies that recognize the H_4 polypeptide. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments, and an Fab expression library. Such an antibody is preferfably specific for human H_4 and it may recognize either a mutant form of H_4 or wild-type H_4, or both.

One can use the hydropathic index of amino acids, as discussed by Kyte and Doolittle (J Mol Biol. 1982, 157:105-132). See, for example, U.S. Patent 4,554,101, which states that the greatest local average hydrophilicity of a "protein," as governed by the hydrophilicity of its adjacent amino acids, correlates with its immunogenicity. Accordingly, it is noted that substitutions can be made based on the hydrophilicity assigned to each amino acid. In using either the hydrophilicity index or hydropathic index, which assigns values to each amino acid, it is preferred to introduce substitutions of amino acids where these values are ± 2, with ± 1 being particularly preferred, and those within ± 0.5 being the most preferred substitutions.

Various procedures known in the art may be used for the production of polyclonal antibodies to H_4 polypeptide or derivative or analog thereof. For the production of antibody, various host animals can be immunized by injection with the H_4 polypeptide, or a derivative (*e.g.*, fragment or fusion protein) thereof, including but not limited to rabbits, mice, rats, sheep, goats, etc. In one embodiment, the H_4 polypeptide or fragment thereof can be conjugated to an

immunogenic carrier, *e.g.*, bovine serum albumin (BSA) or keyhole limpet hemocyanin (KLH). Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and Corynebacterium parvum.

For preparation of monoclonal antibodies directed toward the H_4 polypeptide, or fragment, analog, or derivative thereof, any technique that provides for the production of antibody molecules by continuous cell lines in culture may be used. These include but are not limited to the hybridoma technique originally developed by Kohler and Milstein (Nature 1975, 256:495-497), as well as the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, Immunology Today 1983, 4:72; Cote *et al.*, Proc. Natl. Acad. Sci. 1983, 80:2026-2030), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole *et al.*, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., 1985, pp. 77-96). In an additional embodiment of the invention, monoclonal antibodies can be produced in germ-free animals (International Patent Publication No. WO 89/12690). In fact, according to the invention, techniques developed for the production of "chimeric antibodies" (Morrison *et al.*, J. Bacteriol. 1984, 159:870; Neuberger *et al.*, Nature 1984, 312:604-608; Takeda *et al.*, Nature 1985, 314:452-454) by splicing the genes from a mouse antibody molecule specific for an H_4 polypeptide together with genes from a human antibody molecule of appropriate biological activity can be used; such antibodies are within the scope of this invention. Such human or humanized chimeric antibodies are preferred for use in therapy of human diseases or disorders (described *infra*), since the human or humanized antibodies are much less likely than xenogenic antibodies to induce an immune response, in particular an allergic response, themselves.

Antibody fragments which contain the idiotype of the antibody molecule can be generated by known techniques. For example, such fragments include but are not limited to: the F(ab')2 fragment which can be produced by pepsin digestion of the antibody molecule; the Fab' fragments which can be generated by reducing the disulfide bridges of the F(ab')2 fragment, and the Fab fragments which can be generated by treating the antibody molecule with papain and a reducing agent.

According to the invention, techniques described for the production of single

chain antibodies (U.S. Patent Nos. 5,476,786, 5,132,405, and U.S. Patent 4,946,778) can be adapted to produce H_4 polypeptide-specific single chain antibodies. An additional embodiment of the invention utilizes the techniques described for the construction of Fab expression libraries (Huse *et al.*, Science 1989, 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity for an H_4 polypeptide, or its derivatives, or analogs.

In the production and use of antibodies, screening for or testing with the desired antibody can be accomplished by techniques known in the art, *e.g.*, radioimmunoassay, ELISA (enzyme-linked immunosorbant assay), "sandwich" immunoassays, immunoradiometric assays, gel diffusion precipitin reactions, immunodiffusion assays, *in situ* immunoassays (using colloidal gold, enzyme or radioisotope labels, for example), Western blots, precipitation reactions, agglutination assays (*e.g.*, gel agglutination assays, hemagglutination assays), complement fixation assays, immunofluorescence assays, protein A assays, and immunoelectrophoresis assays, etc. In one embodiment, antibody binding is detected by detecting a label on the primary antibody. In another embodiment, the primary antibody is detected by detecting binding of a secondary antibody or reagent to the primary antibody. In a further embodiment, the secondary antibody is labeled. Many means are known in the art for detecting binding in an immunoassay and are within the scope of the present invention. For example, to select antibodies which recognize a specific epitope of an H_4 polypeptide, one may assay generated hybridomas for a product which binds to an H_4 polypeptide fragment containing such epitope. For selection of an antibody specific to an H_4 polypeptide from a particular species of animal, one can select on the basis of positive binding with H_4 polypeptide expressed by or isolated from cells of that species of animal.

The foregoing antibodies can be used in methods known in the art relating to the localization and activity of the H_4 polypeptide, *e.g.*, for Western blotting, imaging H_4 polypeptide *in situ*, measuring levels thereof in appropriate physiological samples, etc. using any of the detection techniques mentioned above or known in the art. Such antibodies can also be used in assays for ligand binding, *e.g.*, as described in U.S. Patent No. 5,679,582. Antibody binding generally occurs most readily under physiological conditions, *e.g.*, pH of between about 7 and 8, and physiological ionic strength. The presence of a carrier protein in the buffer solutions stabilizes the assays. While there is some tolerance of perturbation of optimal conditions, *e.g.*, increasing or decreasing ionic strength, temperature, or pH, or adding detergents or chaotropic

salts, such perturbations will decrease binding stability.

In a specific embodiment, antibodies that act as ligands and agonize or antagonize the activity of H_4 polypeptide can be generated. In addition, intracellular single chain Fv antibodies can be used to regulate cAMP formation (Marasco et al., Proc. Natl. Acad. Sci. U.S.A. 1993, 90:7884-7893; Chen., Mol. Med. Today 1997, 3:160-167; Spitz et al., Anticancer Res. 1996, 16:3415-22; Indolfi et al., Nat. Med. 1996, 2:634-635; Kijma et al., Pharmacol. Ther. 1995, 68:247-267). Such antibodies can be tested using the assays described *infra* for identifying ligands.

Screening and Chemistry

According to the present invention, nucleotide sequences encoding H_4 and the H_4 receptor structure, which can be modeled from the amino acid sequence based on homology to other GPCR proteins, are useful targets to identify drugs that are effective in treating disorders associated with histamine-regulated processes. Drug targets include without limitation (i) isolated nucleic acids derived from the gene encoding H_4 (*e.g.*, antisense or ribozyme molecules) and (ii)small molecule compounds that recognize and bind the receptor.

In particular, identification and isolation of H_4 provides for development of screening assays, particularly for high throughput screening of molecules that up- or down-regulate the activity of H_4. Accordingly, the present invention contemplates methods for identifying specific histamine receptor ligands that interact with H_4 receptors, using various screening assays known in the art.

Any screening technique known in the art can be used to screen for H_4 agonists or antagonists. The present invention contemplates screens for small molecule ligands or ligand analogs and mimics, as well as screens for natural ligands that bind to and agonize or antagonize H_4 activity *in vivo*. For example, natural products libraries can be screened using assays of the invention for molecules that agonize or antagonize H_4 expression or activity.

Another approach uses recombinant bacteriophage to produce large libraries. Using the "phage method" (Scott and Smith, Science 1990, 249:386-390; Cwirla, *et al.*, Proc. Natl. Acad. Sci., USA 1990, 87:6378-6382; Devlin *et al.*, Science 1990, 49:404-406), very large libraries can be constructed (106-108 chemical entities). A second approach uses primarily chemical methods, of which the Geysen method (Geysen *et al.*, Molecular Immunology 1986,

23:709-715; Geysen *et al.* J. Immunologic Method 1987 102:259-274; and the method of Fodor *et al.* (Science 1991, 251:767-773) are examples. Furka *et al.* (14th International Congress of Biochemistry, Volume #5 1988, Abstract FR:013; Furka, Int. J. Peptide Protein Res. 1991, 37:487-493), Houghton (U.S. Patent No. 4,631,211) and Rutter (U.S. Patent No. 5,010,175) describe methods to produce a mixture of peptides that can be tested as agonists or antagonists.

In another aspect, synthetic libraries (Needels *et al.*, Proc. Natl. Acad. Sci. USA 1993, 90:10700-4; Ohlmeyer *et al.*, Proc. Natl. Acad. Sci. USA 1993, 90:10922-10926; Lam *et al.*, PCT Publication No. WO 92/00252; Kocis *et al.*, PCT Publication No. WO 9428028) and the like can be used to screen for ligands that regulate H_4 activity. Test compounds are screened from large libraries of synthetic or natural compounds. Numerous means are currently used for random and directed synthesis of saccharide, peptide, and nucleic acid based compounds. Synthetic compound libraries are commercially available from Maybridge Chemical Co. (Trevillet, Cornwall, UK), Comgenex (Princeton, NJ), Brandon Associates (Merrimack, NH), and Microsource (New Milford, CT). A rare chemical library is available from Aldrich (Milwaukee, WI). Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available from *e.g.* Pan Laboratories (Bothell, WA) or MycoSearch (NC), or are readily producible. Additionally, natural and synthetically produced libraries and compounds are readily modified through conventional chemical, physical, and biochemical means (Blondelle *et al.*, Tib Tech 1996, 14:60).

Knowledge of the primary sequence of H_4, and the similarity of that sequence with proteins of known function, can provide an initial clue as to the structure of agonists or antagonists of the receptor. Identification and screening of agonists antagonists is further facilitated by determining structural features of the receptor, *e.g.*, using X-ray crystallography, neutron diffraction, nuclear magnetic resonance spectrometry, homology studies, structure-activity relationships, and other techniques for structure determination. These techniques provide for the rational design or identification of agonists and antagonists.

One technique that may be used to assess the affinity of a test compound for the H_4 receptor is a competition binding assay. In this assay, test wells containing an aliquot of a lipid bilayer membranes that contain the histamine H_4 receptor are incubated with an known concentration of a radiolabeled ligand for the receptor. The lipid bilayer may be prepared by any known protocol that separates the membrane containing receptor component from the

cytoplasmic components. Each well also is incubated with a different concentration of a unlabeled test compound. Cell membranes are then separated from the incubation mixture by any method known in the art including, but not limited to, centrifugation and vacuum filtration on a cell harvester. The radioactivity of each well is then determined using any device that can detect radioactivity, such as a scintillation counter. As increasing concentrations of the test compound compete for the receptor binding site, the radioactivity detected decreases. The data then can be converted using the Cheng-Prusoff equation (Biochem Pharmacol. 1973, 22:3099-3108) to determine the affinity (K_i) of the compound for the receptor.

In vivo screening methods

Intact cells or whole animals expressing a gene encoding H_4 can be used in screening methods to identify candidate drugs.

In one series of embodiments, a permanent cell line is established. Alternatively, cells (including without limitation mammalian, insect, yeast, or bacterial cells) are transiently programmed to express an H_4 gene by introduction of appropriate DNA or mRNA. Identification of candidate compounds can be achieved using any suitable assay, including without limitation (i) assays that measure binding of test compounds to H_4 (ii) assays that measure the ability of a test compound to modify (*i.e.*, inhibit or enhance) a measurable activity or function of H_4 and (iii) assays that measure the ability of a compound to modify (*i.e.*, inhibit or enhance) the transcriptional activity of sequences derived from the promoter (*i.e.*, regulatory) regions of the H_4 gene.

H_4 knockout mammals can be prepared for evaluating the molecular pathology of this defect in greater detail than is possible with human subjects. Such animals also provide excellent models for screening drug candidates. A "knockout mammal" is an mammal (*e.g.*, mouse, rabbit) that contains within its genome a specific gene that has been inactivated by the method of gene targeting (see, *e.g.*, U.S. Patent Nos. 5,777,195 and 5,616,491). A knockout mammal includes both a heterozygote knockout (*i.e.*, one defective allele and one wild-type allele) and a homozygous mutant (*i.e.*, two defective alleles; a heterologous construct for expression of an H_4, such as a human H_4, could be inserted to permit the knockout mammal to live if lack of H_4 expression was lethal). Preparation of a knockout mammal requires first introducing a nucleic acid construct that will be used to suppress expression of a particular gene into an undifferentiated cell type termed an embryonic stem cell. This cell is then injected into

a mammalian embryo. A mammalian embryo with an integrated cell is then implanted into a foster mother for the duration of gestation. Zhou, *et al.* (Genes and Development 1995, 9:2623-34) describes PPCA knock-out mice.

The term "knockout" refers to partial or complete suppression of the expression of at least a portion of a protein encoded by an endogenous DNA sequence in a cell. The term "knockout construct" refers to a nucleic acid sequence that is designed to decrease or suppress expression of a protein encoded by endogenous DNA sequences in a cell. The nucleic acid sequence used as the knockout construct is typically comprised of (1) DNA from some portion of the gene (exon sequence, intron sequence, and/or promoter sequence) to be suppressed and (2) a marker sequence used to detect the presence of the knockout construct in the cell. The knockout construct is inserted into a cell, and integrates with the genomic DNA of the cell in such a position so as to prevent or interrupt transcription of the native DNA sequence. Such insertion usually occurs by homologous recombination (*i.e.,* regions of the knockout construct that are homologous to endogenous DNA sequences hybridize to each other when the knockout construct is inserted into the cell and recombine so that the knockout construct is incorporated into the corresponding position of the endogenous DNA). The knockout construct nucleic acid sequence may comprise (1) a full or partial sequence of one or more exons and/or introns of the gene to be suppressed, (2) a full or partial promoter sequence of the gene to be suppressed, or (3) combinations thereof. Typically, the knockout construct is inserted into an embryonic stem cell (ES cell) and is integrated into the ES cell genomic DNA, usually by the process of homologous recombination. This ES cell is then injected into, and integrates with, the developing embryo.

The phrases "disruption of the gene" and "gene disruption" refer to insertion of a nucleic acid sequence into one region of the native DNA sequence (usually one or more exons) and/or the promoter region of a gene so as to decrease or prevent expression of that gene in the cell as compared to the wild-type or naturally occurring sequence of the gene. By way of example, a nucleic acid construct can be prepared containing a DNA sequence encoding an antibiotic resistance gene which is inserted into the DNA sequence that is complementary to the DNA sequence (promoter and/or coding region) to be disrupted. When this nucleic acid construct is then transfected into a cell, the construct will integrate into the genomic DNA. Thus, many progeny of the cell will no longer express the gene at least in some cells, or will express it at a

decreased level, as the DNA is now disrupted by the antibiotic resistance gene.

Generally, the DNA will be at least about 1 kb in length and preferably 3-4 kb in length, thereby providing sufficient complementary sequence for recombination when the knockout construct is introduced into the genomic DNA of the ES cell (discussed below).

Included within the scope of this invention is a mammal in which two or more genes have been knocked out. Such mammals can be generated by repeating the procedures set forth herein for generating each knockout construct, or by breeding to mammals, each with a single gene knocked out, to each other, and screening for those with the double knockout genotype.

Regulated knockout animals can be prepared using various systems, such as the tet-repressor system (see U.S. Patent No. 5,654,168) or the Cre-Lox system (see U.S. Patent Nos. 4,959,317 and 5,801,030).

In another series of embodiments, transgenic animals are created in which (i) a human H_4 is stably inserted into the genome of the transgenic animal; and/or (ii) the endogenous H_4 genes are inactivated and replaced with human H_4 genes. See, *e.g.*, Coffman, Semin. Nephrol. 1997, 17:404; Esther *et al.*, Lab. Invest. 1996, 74:953; Murakami et al., Blood Press. Suppl. 1996, 2:36.

H_4 Activation Assay

Any cell assay system that allows for assessment of functional activity of H_4 agonists and antagonists is defined by the present invention. In a specific embodiment, exemplified *infra*, the assay can be used to identify compounds that selectively interact with H_4, which can be evaluated by assessing the effects of H_4 transformed cells contacted with a test compound, which modulates cAMP accumulation. The assay system can thus be used to identify compounds that selectively produce a functional effect through histamine H_4 receptors. Compounds that increase cAMP formation and accumulation may be useful as novel therapeutics in the prevention of transplanted organ rejection, asthma, allergies, and autoimmune pathologies, such as, but not limited to, multiple sclerosis, type I diabetes, and rheumatoid arthritis. Additionally, CNS functions such as cognitive and memory defects may be treated. Preferably, each experiment is performed in triplicate at multiple different dilutions of test compound.

An agonist and/or antagonist screen involves detecting cAMP accumulation by the host cell when contacted with H_4 ligand. If cAMP accumulation is increased, the test

compound is a candidate antagonist of H_4 receptors. If cAMP accumulation is decreased, the test compound is a candidate agonist of H_4 receptors. If there is no change in cAMP formation, the test compound is not an effective H_4 ligand.

Any convenient method permits detection of the formed product, cAMP. For example, the invention provides immunoassays for detecting cAMP. Typically, immunoassays use either a labeled antibody or a labeled antigenic component (*e.g.*, that competes with the antigen in the sample for binding to the antibody). Suitable labels include without limitation enzyme-based, fluorescent, chemiluminescent, radioactive, or dye molecules. Assays that amplify the signals from the probe are also known, such as, for example, those that utilize biotin and avidin, and enzyme-labelled immunoassays, such as ELISA assays. Alternatively, labeled antigenic component may be quantified by scintillation techniques. In another method, the second messenger, preferably cAMP, will be separated on a high performance liquid chromatograph and quantified by a UV detector.

The assay system described here also may be used in a high-throughput primary screen for agonists and antagonists, or it may be used as a secondary functional screen for candidate compounds identified by a different primary screen, *e.g.*, a binding assay screen that identifies compounds that interact with the receptor.

High-Throughput Screen

Agents according to the invention may be identified by screening in high-throughput assays, including without limitation cell-based or cell-free assays. It will be appreciated by those skilled in the art that different types of assays can be used to detect different types of agents. Several methods of automated assays have been developed in recent years so as to permit screening of tens of thousands of compounds in a short period of time. Such high-throughput screening methods are particularly preferred. The use of high-throughput screening assays to test for agents is greatly facilitated by the availability of large amounts of purified polypeptides, as provided by the invention.

Compounds

"Histamine" refers to a neurotransmitter that is produced and released from neurons. Histamine is formed from the amino acid histidine by histidine decarboxylase. Structurally, histamine is an imidazolethylamine. In other words, histamine is comprised of

an imidazole moiety and an amino group connected by an alkyl chain. The pharmacologically active form of histamine is proposed to be the monocationic tautomer, where one of the nitrogen present in the imidazole ring is positively charged. However, different forms of histamine may interact with histamine receptors to produce a functional effect. Histamine is produced intracellularly and stored until released in response to a physiological stimulus.

"Histamine analogs" refers to compounds that comprise the imidazole, alkyl chain, and a nitrogen moiety (*e.g.*, amine (unsubstituted or substituted), piperidine, pyridine) of histamine, but may be modified at other positions. These modifications may be performed to alter affinity and/or selectivity of the compound for the histamine receptors. "Histamine compounds" refers to compounds that may bind to the histamine receptors.

Histamine analogs and compounds can be classified as agonists or antagonists. As discussed previously, agonists are ligands that bind to the receptor and produce a functional effect similar to that produced by the endogenous ligand (*i.e.*, histamine) for the receptor, whereas antagonists are ligands that bind to the receptor and block a functional effect produced by an agonist for the receptor or the endogenous ligand of the receptor. Histamine analogs and compounds are further described in Goodman and Gilman's The Pharmacological Basis of Therapeutics, Ninth Edition, McGraw-Hill, 1996.

Agonists that may be contemplated by this invention include, but are not limited to, R-(α)-methylhistamine, imetit, and immepip. Antagonists, burimamide, impromidine, dimaprit, and thioperamide clobenpropit and iodophenpropit impentamine, GT2016 and iodoproxyfan. Other compounds include derivatives, metabolites, and precursors.

Methods of Diagnosis

According to the present invention, genetic variants of H_4 can be detected to diagnose an H_4 associated disease, such as treatment for transplanted organ rejection, asthma, allergies, and autoimmune pathologies, such as, but not limited to, multiple sclerosis, type I diabetes, and rheumatoid arthritis. Additionally, CNS functions such as cognitive and memory defects may be treated. The various methods for detecting such variants are described herein. Where such variants impact H_4 function, either as a result of a mutated

amino acid sequence or because the mutation results in expression of a truncated protein, or no expression at all, they are expected to result in disregulation of the allergic response, the immune response, cognition and memory.

Nucleic Acid Assays

The DNA may be obtained from any cell source. DNA is extracted from the cell source or body fluid using any of the numerous methods that are standard in the art. It will be understood that the particular method used to extract DNA will depend on the nature of the source. Generally, the minimum amount of DNA to be extracted for use in the present invention is about 25 pg (corresponding to about 5 cell equivalents of a genome size of 4 x 10^9 base pairs).

In another alternate embodiment, RNA is isolated from biopsy tissue using standard methods well known to those of ordinary skill in the art such as guanidium thiocyanate-phenol-chloroform extraction (Chomocyznski *et al.*, Anal. Biochem., 162:156, 1987). The isolated RNA is then subjected to coupled reverse transcription and amplification by polymerase chain reaction (RT-PCR), using specific oligonucleotide primers that are specific for a selected site. Conditions for primer annealing are chosen to ensure specific reverse transcription and amplification; thus, the appearance of an amplification product is diagnostic of the presence of a particular genetic variation. In another embodiment, RNA is reverse-transcribed and amplified, after which the amplified sequences are identified by, *e.g.*, direct sequencing. In still another embodiment, cDNA obtained from the RNA can be cloned and sequenced to identify a mutation.

Protein Assays

In an alternate embodiment, biopsy tissue is obtained from a subject. Antibodies that are capable of specifically binding to H_4 are then contacted with samples of the tissue to determine the presence or absence of a H_4 polypeptide specified by the antibody. The antibodies may be polyclonal or monoclonal, preferably monoclonal. Measurement of specific antibody binding to cells may be accomplished by any known method, *e.g.*, quantitative flow cytometry, enzyme-linked or fluorescence-linked immunoassay, Western analysis, etc.

Therapeutic Uses

According to the present invention, stimulation of H_4 receptor activity may be used as a treatment option in patients with histamine-related disease states. Stimulation of H_4 receptor activity may be by methods, such as, but not limited to, (i) providing polypeptides that stimulate receptor activity and (ii) providing compounds that stimulate receptor activity.

Gene Therapy

In a specific embodiment, vectors comprising a sequence encoding a protein, including, but not limited to, full-length H_4, are provided to treat or prevent a disease or disorder associated with the function of H_4 in peripheral blood leukocytes. In this embodiment of the invention, the therapeutic vector encodes a sequence that produces the protein of the invention.

Any of the methods for gene therapy available in the art can be used according to the present invention. Exemplary methods are described below.

For general reviews of the methods of gene therapy, see, Goldspiel *et al.*, Clinical Pharmacy, 1993, 12:488-505; Wu and Wu, Biotherapy, 1991, 3:87-95; Tolstoshev, Ann. Rev. Pharmacol. Toxicol., 1993, 32:573-596; Mulligan, Science, 1993, 260:926-932; and Morgan and Anderson, Ann. Rev. Biochem., 1993, 62:191-217; May, TIBTECH, 1993, 11:155-215. Methods commonly known in the art of recombinant DNA technology that can be used are described in Ausubel *et al.*, (eds.), 1993, Current Protocols in Molecular Biology, John Wiley & Sons, NY; Kriegler, 1990, Gene Transfer and Expression, A Laboratory Manual, Stockton Press, NY; and in Chapters 12 and 13, Dracopoli et al., (eds.), 1994, Current Protocols in Human Genetics, John Wiley & Sons, NY. Vectors suitable for gene therapy are described above.

In one aspect, the therapeutic vector comprises a nucleic acid that expresses a protein of the invention in a suitable host. In particular, such a vector has a promoter operationally linked to the coding sequence for the protein. The promoter can be inducible or constitutive and, optionally, tissue-specific. In another embodiment, a nucleic acid molecule is used in which the protein coding sequences and any other desired sequences are flanked by regions that promote homologous recombination at a desired site in the genome, thus providing for intrachromosomal expression of the protein (Koller and Smithies, Proc. Natl. Acad. Sci. U.S.A, 1989, 86:8932-8935; Zijlstra *et al.*, Nature, 1989, 342:435-438).

Delivery of the vector into a patient may be either direct, in which case the patient is directly exposed to the vector or a delivery complex, or indirect, in which case, cells are first

transformed with the vector *in vitro* then transplanted into the patient. These two approaches are known, respectively, as *in vivo* and *ex vivo* gene therapy.

In a specific embodiment, the vector is directly provided *in vivo*, where it enters the cells of the organism and mediates expression of the protein. This can be accomplished by any of numerous methods known in the art, by constructing it as part of an appropriate expression vector and administering it so that it becomes intracellular, *e.g.*, by infection using a defective or attenuated retroviral or other viral vector (see, U.S. Patent No. 4,980,286), or by direct injection of naked DNA, or by use of microparticle bombardment (*e.g.*, a gene gun; Biolistic, Dupont); or coating with lipids or cell-surface receptors or transfecting agents, encapsulation in biopolymers (*e.g.*, poly-S-1-64-N-acetylglucosamine polysaccharide; see, U.S. Patent No. 5,635,493), encapsulation in liposomes, microparticles, or microcapsules; by administering it in linkage to a peptide or other ligand known to enter the nucleus; or by administering it in linkage to a ligand subject to receptor-mediated endocytosis (see, *e.g.*, Wu and Wu, J. Biol. Chem., 1987, 62:4429-4432), etc. In another embodiment, a nucleic acid ligand complex can be formed in which the ligand comprises a fusogenic viral peptide to disrupt endosomes, allowing the nucleic acid to avoid lysosomal degradation. In yet another embodiment, the nucleic acid can be targeted *in vivo* for cell specific uptake and expression, by targeting a specific receptor (see, *e.g.*, PCT Publication Nos. WO 92/06180, WO 92/22635, WO 92/20316 and WO 93/14188). Alternatively, the nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression by homologous recombination (Koller and Smithies, Proc. Natl. Acad. Sci. USA, 1989, 86:8932-8935; Zijlstra, et al., Nature, 1989, 342:435-438). These methods are in addition to those discussed above in conjunction with "Viral and Non-viral Vectors".

The form and amount of therapeutic nucleic acid envisioned for use depends on the type of disease and the severity of the desired effect, patient state, etc., and can be determined by one skilled in the art.

Inhibition or stimulation of protein synthesis

Gene transcription and protein translation may be inhibited or stimulated by administration of exogenous compounds. Exogenous compounds may interact with extracellular and/or intracellular messenger systems, such as, but not limited to, adenosine triphosphate, nitric oxide, and guanosine triphosphate; to regulate protein synthesis. In this embodiment, exogenous compounds that stimulate or inhibit H_4 protein synthesis may be

used in the prevention and/or treatment for transplanted organ rejection, asthma, allergies, and autoimmune pathologies, such as, but not limited to, multiple sclerosis, type I diabetes, and rheumatoid arthritis. Additionally, CNS functions such as cognitive and memory defects may be treated.

The present invention provides antisense nucleic acids (including ribozymes), which may be used to inhibit expression of H_4 of the invention. The antisense nucleic acid, upon hybridizing under cytoplasmic conditions with complementary bases in an RNA or DNA molecule, inhibits the role of the RNA or DNA. Additionally, hybridization of the antisense nucleic acid to the DNA or RNA may inhibit transcription of the DNA into RNA and/or translation of the RNA into the protein. If the RNA is a messenger RNA transcript, the antisense nucleic acid is a countertranscript or mRNA-interfering complementary nucleic acid. Antisense nucleic acid molecules can be encoded by a recombinant gene for expression in a cell (*e.g.*, U.S. Patent No. 5,814,500; U.S. Patent No. 5,811,234) or can be prepared synthetically (*e.g.*, U.S. Patent No. 5,780,607).

Alternatively, antibody molecules can also be administered, for example, by expressing nucleotide sequences encoding single-chain antibodies within the target cell population by utilizing, for example, techniques such as those described in Marasco *et al.* (Proc. Natl. Acad Sci. USA, 1993, 90:7889-7893).

Therapeutically suggested compounds may be provided to the patient in formulations that are known in the art and may include any pharmaceutically acceptable additives, such as excipents, lubricants, diluents, flavorants, colorants, and disintegrants. The formulations may be produced in useful dosage units such as tablet, caplet, capsule, liquid, or injection.

The form and amount of therapeutic compound envisioned for use depends on the type of disease and the severity of the desired effect, patient state, etc., and can be determined by one skilled in the art.

EXAMPLES

The present invention will be better understood by reference to the following Examples, which are provided as exemplary of the invention, and not by way of limitation.

EXAMPLE I: CHARACTERIZATION OF THE H_4 RECEPTOR

A GPCR sequence profile was generated from a sequence alignment of members of this subfamily based on hidden Markov Models (Eddy Bioinformatics, 1998 14:755-763; Durbin *et al.* A tutorial introduction to hidden Markov models and other probabilistic modeling approaches in computational sequence analysis, Cambridge University Press, 1998) to look for novel members in the human genome database. In detail, GPCR proteins of the biogenic amine subfamily were retrieved from Swiss-Prot database and a sequence alignment was generated by a multiple sequence alignment tool, named CLUSTALW (Thompson *et al.* Nucleic Acids Res, 1994, 22:4673-80). Using the HMMER program (Eddy, HMMER User's Guide and Program, Version 2.1, 1998), a consensus sequence (GNLLVILVIL RTKKLRTPTN IFILNLAVAD LLFLLTLPPW ALYYLVGGSE DWPFGSALCK LVTALDVVNM YASILLLTAI SIDRYLAIVH PLRYRRRRTS PRRAKVVILL VWVLALLLSL PPLLFSWVKT VEEGNGTLNV NVTVCLIDFP EESTASVSTW LRSYVLLSTL VGFLLPLLVI LVCYTRILRT LRKAAKTLLV VVVVFVLCWL PYFIVLLLDT LCLSIIMSST CELERVLPTA LLVTLWLAYV NSCLNPIIY; SEQ ID NO: 3) was developed from the biogenic amine subfamily members. The consensus sequence contained the unique 7-transmembrane sequence structure of biogenic amine GPCRs.

A weekly update of nucleotide sequence from GenBank database is maintained in-house. An auto-search script using TBLASTN program was written and the biogenic amine GPCR consensus sequence was used to search this database weekly. Every TBLASTN search result was carefully examined and potential open reading frame (ORF) fragments were extracted from nucleotide sequence. Each fragment was further verified to determine its novelty.

Four peptide fragments translated from a recently released human genomic sequence (Accession number: AC007922) of chromosome 18 clone RP11-178F10, from the Whitehead Institute/MIT Center for Genome Research, were shown to have modest homology to different regions of this GPCR consensus sequence. These four fragments located to different regions of sequence AC009722 in both plus and minus strands. Interestingly, the highest scoring hit from searching the protein database with these peptides by BLASP was in all cases, the human histamine receptor 3 (H_3). It is very likely that these four fragments are exons of a GPCR gene and their appearance in both strands of this genomic sequence may result from incorrect genomic contig assembly.

To clone the full length ORF, multiple primer sets were designed against the predicted sequence. Using multiple primer pairs, 3 overlapping pieces spanning the entire ORF were obtained from a human heart marathon cDNA library (Clontech, Palo Alto, CA). Using the marathon adapter primer API, and a reverse primer corresponding to ORF bases 644-615, a fragment of the H_4 cDNA was obtained which corresponds to bases 276-644 of the H_4 coding region (exon sequence) and 39 bases of intronic sequence at the 5' end. Primers comprising nucleotides 461-482 and 1173-1146 were used to PCR a 712bp band. The 5' sequence was obtained using primers comprising nucleotides 1-32 and 339-309. The outer primers (1-32, 1173-1146) were used to piece the 3 fragments together. The sequence was ligated into the mammalian expression vector pCDNA3.1 + zeo (Invitrogen, Carlsbad, CA).

EXAMPLE II: TISSUE EXPRESSION OF THE H_4 RECEPTOR

Quantitative RT-PCR was performed on an ABI 7700 "Taqman" sequence detection system to determine the tissue distribution of the H_4 receptor. Primers spanning the exon 1-2 boundary (Forward primer: 5'-taacttggccattgacttctt-3' (SEQ ID NO:4); Reverse primer: 5'-attcgaacagcatgtgagggat-3' (SEQ ID NO:5) and a Probe: 5'-(6-carboxyfluorescein)-tacaaaggaatggagatcacacccaca-(6-carboxy-N,N,N'N'-tetramethylrhodamine)-3' (SEQ ID NO:6)) were used to determine H_4 expression levels in a mRNA prepared from a series of human tissues, purchased from Clontech (Palo Alto). A 2 step reaction procedure was performed as per manufacturers directions. Briefly, 2mg RNA was reverse transcribed using random hexamers (2.5mM in a final volume of 20ml. 14ul of this was used in the PCR reaction. The cyling conditions were as follows, 95 °C for 10 minutes, followed by 40 cycles of (a) 95 °C for 15 seconds and (b) 60 °C for 1 minute.

EXAMPLE III: PHARMACOLOGICAL PROFILE OF THE H_4 RECEPTOR IN YEAST CELLS

The H_4 receptor protein coding sequences were amplified using Forward oligo: 5'-aaggatccaaaatgccagatactaatagc-3' (SEQ ID NO:7) and Reverse oligo: 5'-aagtcgacttaagaagatactgaccgac-3' (SEQ ID NO:8) that add BamHI and yeast consensus translational initiation sites to the 5' end and a SalI site to the 3' end. The fragment was cloned into corresponding sites in the multicopy yeast expression vector, p426GPD, thus placing

receptor expression under control of the strong constitutive GPD1 promoter. The yeast expression plasmid, pMP327, was introduced into MPY578i5 cells (MATa ura3 his3 trp1 leu2 lys2 ade2 far1::LYS2 fus1::FUS1-HIS3 sst2::SST2-G418R ste2::LEU2, gpa1::GPA1i5) (Hadcock, J.R. and Pausch, M.H. submitted) using LiOAc and selected for ura prototrophy. In order to facilitate coupling of the receptor to the G protein, MPY578i5 cells express a chimeric G alpha protein coupled to the mating signal transduction pathway. The chimeric construct is expressed from the GPA1 locus and is composed of Gpa1 sequences in which the 5'-C-terminal amino acids have been replaced with those of Gai3. A multicopy FUS1-LacZ reporter gene plasmid, pMP283 (Hadcock, J.R. and Pausch, M.H. submitted), was subsequently introduced into H_4R-containing MPY578i5 cells and selected on media lacking trp and ura. The resulting yeast strain, MPY733, was used for further analysis.

Samples (250 ng) of compounds present in the LOPAC panel (Sigma RBI, Natick MA). were dispensed to 96 well microtiter dishes. MPY733 cells (5 x 10^5/ml, 200 ml/well) in assay medium (SCD-ura-trp, pH 6.8, 25 mM PIPES, 0.1mg/ml Chlorophenylred bglactopyranoside(CPRG), 2 mM 3-AT) were added and cultured overnight at 30 °C. The presence of active compounds was detected the next day by measurement of absorbance at 570 nm using a Wallac Victor II. The LOPAC panel was screened in duplicate and in parallel with another yeast strain containing a different orphan GPCR. Only compounds that produced significantly elevated absorbance in both receptor containing plates and not in the other GPCR containing plates were deemed active.

EXAMPLE IV: PHARMACOLOGICAL PROFILE OF THE H_4 RECEPTOR IN MAMMALIAN CELLS

The ORF was modified by PCR for mammalian expression of H_4. A 5' HindIII restriciton enzymes site and a Kozak consensus sequence were added using the primer 5'-aagcttccaccatgccagatactaatagcacaatcaatttatc-3' (SEQ ID NO:9), and a 3' Xba1 site added 5'-tctagattaagaagatactgaccgactgtgttg-3' (SEQ ID NO:10). The sequence was confirmed and ligated into the HindIII and Xba I sites of the mammalian expression vector pCDNA3.1+ zeo (Invitrogen, Carlsbad CA). HEK 293 cells (approximately 10^7 cells) were transfected with the pCDNA3.1+zeo/H_4 using standard lipofectamine plus reagent (Life Technologies). Cells were maintained in DMEM containing 10% fetal calf serum and penicillin (100units/ml)/streptomycin

(100mg/ml). 48 hours post transfection cells were selected using 500mg/ml zeocin. Zeocin resistant clones were assayed using the cAMP assay and by RT-PCR. RNA was extracted from approximately 10^6 cells using One step PCR kit (Life technologies). The primers used in the extraction were 5'-ggaaggatgaaggtagtgaatg-3' (SEQ ID NO:11) and 5'-cagaatctgattgggaggaagg-3' (SEQ ID NO:12).

HEK cells stably expressing the H_4 receptor were assessed functionally in cAMP assays using the cAMP scintillation proximity assay (SPA) (Amersham Pharmacia Biotech, Pistcataway, NJ). Briefly, 40 000 cells were plated into wells of a 96 well plate. 24 hours later the media was removed and replaced with 100µl Krebs bicarbonate buffer and the cells were incubated at 37°C for 15 minutes. Following this, the cells were incubated in Krebs buffer containing 0.5mM isobutylmethylxanthine, a phosphodiesterase inhibitor, to prevent cAMP breakdown. The effect of H_4 on the forskolin induced formation was determined by incubating the cells in the presence of forskolin (10µM) and agonist for 12 minutes. cAMP levels were determined using the cAMP SPA kit according to the manufacturers directions.

EXAMPLE V: H_4 EXPRESSION IN INDUCED CD4 AND CD8 CLONES

This examination was undertaken using the TaqMan EZ RT PCR kit and the following oligonucleotides (provided by Philip Jones at Wyeth Neuroscience) (i) H_4 EX 1F (forward oligo): 5'-taacttggccatctttgac-3' (SEQ ID NO:13), (ii) H_4 EX 1R (reverse primer): 5'-attcgaacagcgtgtgag-3' (SEQ ID NO:14), and H_4 EX 1 Probe: 5'-(6-carboxyfluorescein)-tacaaaggaatggagatca-3' (SEQ ID NO:15). RNA for the standard curve was polyA+ human leukocyte RNA from Clontech.

50 nanograms of total RNA was assayed in duplicate from 6 human T cell clones derived from a single human donor. Three of these clones were CD4+, two were CD8+, and one was an NK T cell clone. These T cell clones were stimulated with anti-CD3 and RNA was isolated at 0, 2, 4, 8, 24 and 48 hours after stimulation. Expression was normalized with GAPDH.

In order to determine whether the pattern of H_4 expression was specific to clonal populations of human lymphocytes, the same TaqMan quantitative PCR assay on 10 nanograms of RNA from bulk populations of CD4+ and CD8+ T cells isolated from three different individuals. These lymphocytes were stimulated with anti-CD3 and harvested at the same time

points listed above. H_4 expression was normalized with GAPDH.

RESULTS

Using a GPCR sequence profile generated from a Hidden Markov Model (HMM) of the biogenic amine subfamily, several sequences with homology to the human H3 receptor were identified from the human genomic sequence of chromosome 18 (clone RP11- 178F10 Accession number: AC007922 from the Whitehead Institute/MIT Center for Genome Research) (Eddy Bioinformatics, 1998, 14:755-763). A contig of the predicted exons forms a sequence encoding a putative GPCR whose nucleotide sequence and translated peptide sequence are shown in Figure 1 and SEQ ID NOS:1 and 2. The sequence was deposited with Genbank (Accession Number AF307973).

The predicted protein sequence contains residues that are conserved across the biogenic amine receptors. As in other biogenic amine receptors, the present protein contains conserved aspartate residues within transmembrane (TM) domains 2 and 3 (at positions 61 and 94). A DY (D94 and Y95) motif in TM3 is found only in histamine and muscarinic receptors. A DRY sequence motif that exists at the cytoplasmic interface of TM3 of all GPCRs also can be found in the H_4 sequence (amino acids 111- 113). Also present are two conserved cysteine residues in the first and second extracellular loops (position 88 and 164), which are predicted to form a disulfide bond; two conserved tryptophan residues in TM4 and TM6 (positions 140 and 316); and conserved proline residues in TM5, TM6, and TM7 (positions 186, 318, and 355, respectively). Based on these results, H_4 cDNA appears to encode a biogenic amine-like receptor.

The predicted open reading frame is 1173 base pairs long and encodes a protein of 390 amino acids. Sequence comparison using BLASTP, under standard conditions, analysis reveals that the novel protein is most similar to the human H_3 receptor (44% identical and 51% similar). Based on sequence homology it is proposed that the receptor belongs to the histamine receptor family, therefore we have termed it the H_4 receptor.

Receptor distribution studies indicate that the present receptor is highly expressed in peripheral blood leukocytes. Trace amounts of the H_4 receptor are expressed in heart, lung and placenta. It is proposed that these trace amounts of expression represent blood cell mRNA present in the samples.

Functional studies further confirmed the categorization of the protein as a novel histamine receptor. Histamine and the H_3 selective agonist R-α-methylhistamine stimulated cAMP accumulation, suggesting these compounds are agonists at the H_4 receptor (See Figure 2A). R-α-methylhistamine exhibited both lower potency and efficacy than histamine at the H_4 receptor despite being more potent at the related cloned H_3 receptor and in several tissue based assays for the H_3 receptor (Vollinga, *et al.*, J. Med. Chem. 1995, 38:266-271; Harper, *et al.*, Br. J. Pharmacol. 1999, 128:751-759). The selective H_3 receptor antagonist clobenpropit was a partial agonist at the H_4 receptor (See Figure 2A).

The H_3 antagonist thioperamide almost fully inhibited the stimulatory response produced by histamine (See Figure 2B). Comparatively, clobenpropit partially blocked histamine-induced stimulation of cAMP formation, further suggesting that clobenpropit is a partial agonist at the H_4 receptor. The interaction of the H_4 receptor to a GPAI/ Gαi3 chimeric G-protein alpha subunit also predicts its coupling specificity in mammalian cells.

Expression of the H_4 receptor in HEK 293tsa cells confirms the coupling of the H_4 receptor to the inhibition of cAMP formation (See Figure 2C). In these cells, forskolin stimulated 8-fold greater cAMP formation compared to basal levels. Addition of histamine (1 μM) inhibited forskolin-induced stimulation of cAMP accumulation by about 40%. These studies suggest that the H_4 receptor couples an inhibitory G-protein to inhibit adenylyl cyclase activity and cAMP accumulation.

Structure-activity relationship studies were conducted with several histamine antagonists to further define the pharmacological profile of this receptor. Studies were conducted to correlate the effect of the alkyl chain to antagonist activity at the histamine H_4 receptor (See Figure 3). The rank order of efficacy obtained for the human H_4 was (number represents the length of the alkyl chain) 5>6= 1 0>8=4. This rank order of efficacy contrasts with the human H_3 receptor where the rank order is 5>4>6>8>10.

Prior studies have indicated the expression of at least two subtypes of the H_3 receptor. These two potential H_3 subtypes, localized in the rat brain and guinea pig jejenum, have been shown to have rank order of efficiencies of 4=5>3>6>8 and 5>6=4>8>3, respectively. These assays were different (rat was a radioligand binding assay and guinea pig was an organ bath experiment using isolated guinea pig ileum). However, the relative potencies of the series can be compared, so its possible to say that as the rank order differs then the receptors are likely

to be different (Leurs *et al.* J. Pharm. Exp. Ther., 1996, 276:1009-1015.)

The present studies also indicate that thioperamide is more efficacious and potent than impentamine at the H_4 receptor. Comparatively, impentamine has greater affinity for cloned histamine H_3 receptors (4-fold) than thioperamide (Ki=50.8nM and 193nM respectively).

H_4 expression was consistently detected in all three CD4+ clones at most time points (Figure 4). There was no consistent temporal pattern to anti-CD3 induction of expression between these three clones. One CD8+ clone showed a very low but detectable level of expression at most time points, while the other as well as the NKT clone had no detectable H_4 expression despite good GAPDH amplification.

H_4 expression, normalized with GAPDH, was detectable in all samples at all time points (Figure 5). Anti-CD3 stimulation produced significant induction of H_4 expression in all 6 samples, most peaking at the 8 hour time point. The highest level of induction was found in the 8 hour CD4+ and CD8+ samples from the same donor (designated 'U'), peaking at greater than 100 copies of H_4 RNA per copy of GAPDH.

Combined, these studies indicate that the pharmacological profile of the H_4 receptor is not similar to any known histamine H_3 receptor.

* * *

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and the accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

It is further to be understood that values are approximate, and are provided for description.

Patents, patent applications, publications, procedures, and the like are cited throughout this application, the disclosures of which are incorporated herein by reference in their entireties.

WHAT IS CLAIMED IS:

1. An isolated H_4 receptor protein having an amino acid sequence at least 51% identical or comprising at least 10 contiguous amino acids from the sequence depicted in SEQ ID NO:2 which H_4 receptor protein binds ligands comprising an imidazole and an amine, which imidazole and amine are attached by an alkyl chain, wherein the rank order of efficacy of modulation second messenger formation of the ligands at the H_4 receptor protein is 5>6=10>8=4, where the number represents the number of carbons in the alkyl chain.

2. The H_4 receptor protein of claim 1, wherein upon binding histamine or a histamine agonist the receptor protein inhibits second messenger formation.

3. The H_4 receptor protein of claim 2, wherein the second messenger is cAMP.

4. The H_4 receptor protein of claim 1 which is a human H_4 receptor protein.

5. The H_4 receptor protein of claim 4 which has an amino acid sequence as depicted in SEQ ID NO: 2.

6. The H_4 receptor protein of claim 4 which is encoded by a nucleic acid having a sequence as depicted in SEQ ID NO: 1.

7. An isolated H_4 receptor protein having an amino acid sequence with at least 95% sequence identity to human H_4 receptor protein having an amino acid sequence as depicted in SEQ ID NO: 2.

8. An isolated nucleic acid encoding the H_4 receptor protein of claim 1 or 7.

9. The nucleic acid of claim 8 which is a cDNA.

10. The nucleic acid of claim 8, wherein the H_4 receptor protein is a human H_4 receptor protein.

11. An isolated nucleic acid encoding an H_4 receptor protein, which nucleic acid hybridizes under stringent conditions to a nucleic acid having a sequence of at least 20 nucleotides identical to a corresponding nucleotide sequence of the same number of bases in SEQ ID NO:1 or its complement.

12. The nucleic acid of claim 8, which encodes an H_4 receptor protein having an amino acid sequence as depicted in SEQ ID NO:2.

13. The nucleic acid of claim 12, which comprises a nucleotide sequence as depicted in SEQ ID NO:1.

14. A vector comprising the nucleic acid of claim 8 operably associated with an expression control sequence.

15. A host cell transfected with the vector of claim 14.

16. A non-human animal transformed with the vector of claim 14, wherein the animal expresses a H_4 receptor protein at a detectable level, whereby the cells expressing the H_4 receptor protein suppress cAMP formation when contacted with an H_4 receptor agonist.

17. A method for producing a H_4 receptor protein, which method comprises culturing host cells of claim 15 under conditions that provide for expression of the H_4 receptor protein by the vector.

18. An isolated nucleic acid of at least ten bases having a nucleotide sequence identical to a corresponding nucleotide sequence of the same number of bases in SEQ ID NO: 1 or its complement.

19. The nucleic acid of claim 18 which is detectably labeled.

20. An antibody that specifically binds to the H_4 receptor protein of claim 1 or 7.

21. A method for detecting an H_4 receptor protein, which method comprises detecting binding of the antibody of claim 20 to a protein in a sample suspected of containing a H_4 receptor protein, wherein the antibody is contacted with the sample under conditions that permit specific binding with any H_4 receptor protein present in the sample.

22. A method for detecting expression of H_4 receptor, which method comprises detecting mRNA encoding H_4 receptor in a sample from a cell suspected of expressing H_4 receptor.

23. The method according to claim 22 wherein mRNA encoding H_4 receptor is detected by hybridization to a H_4 receptor-specific nucleic acid.

24. The method according to claim 23 wherein the H_4 receptor-specific nucleic acid is at least 10 nucleotides in length and has a sequence identical to a sequence of the same number of bases in SEQ ID NO: 1, or the complementary sequence thereof.

25. An assay system for identifying H_4 receptor ligands, comprising a sufficient number of cells of claim 15 to detect an alteration in second messenger accumulation.

26. The assay system of claim 25, wherein the second messenger is cAMP.

27. The assay system of claim 25, wherein the receptor is a human receptor.

28. A method for identifying a test compound that antagonizes histamine H_4 receptors, which method comprises detecting an increase in the level of a second messenger in an assay system of claim 25 contacted with the test compound, wherein an increase in the level of the second messenger indicates that the test compound antagonizes the H_4 receptor.

29. A method for identifying a test compound that agonizes histamine H_4 receptors, which method comprises detecting a decrease in the level of a second messenger in an assay system of claim 25 contacted with the test compound, wherein the decrease in the level of the second messenger indicates that the test compound agonizes the H_4 receptor.

30. A method for identifying a compound that binds an H_4 receptor, which method comprises detecting binding of a test compound to the H_4 receptor protein of claim 1.

31. The method according to claim 29, wherein binding of the test compound is detecting by inhibiting binding of a labeled H_4 ligand.

32. The method according to claim 29, wherein the H_4 receptor protein is in a lipid bilayer membrane.

33. An isolated nucleic acid that specifically hybridizes under highly stringent conditions to the complement of the sequence depicted in SEQ ID NO:1, wherein said nucleic acid encodes a histamine H_4 receptor protein.

Figure 1

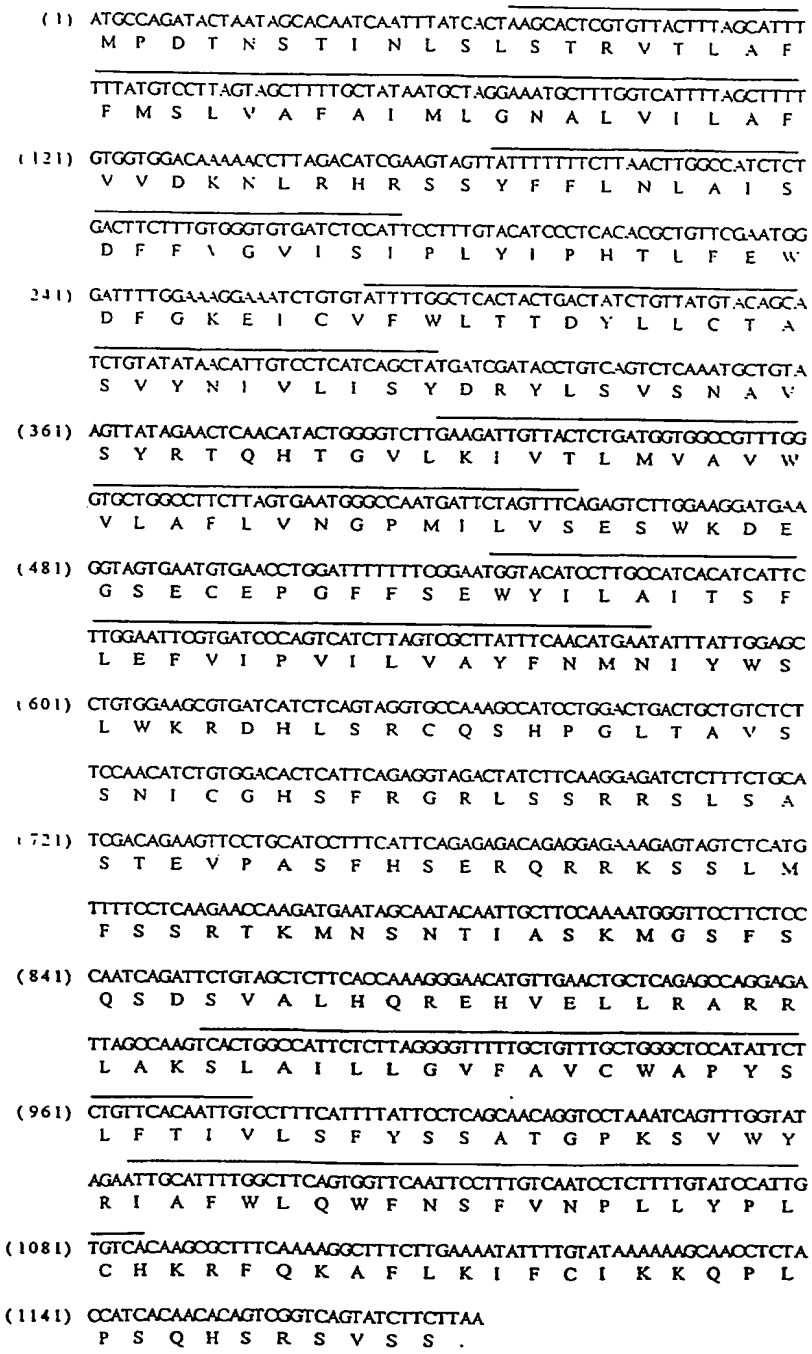
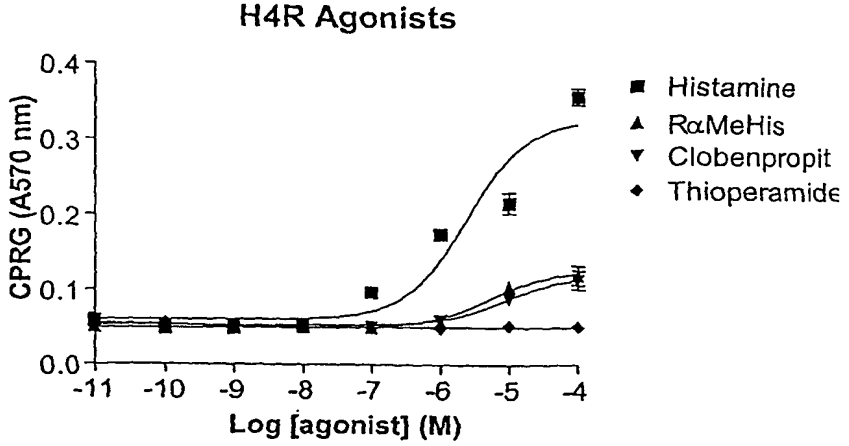
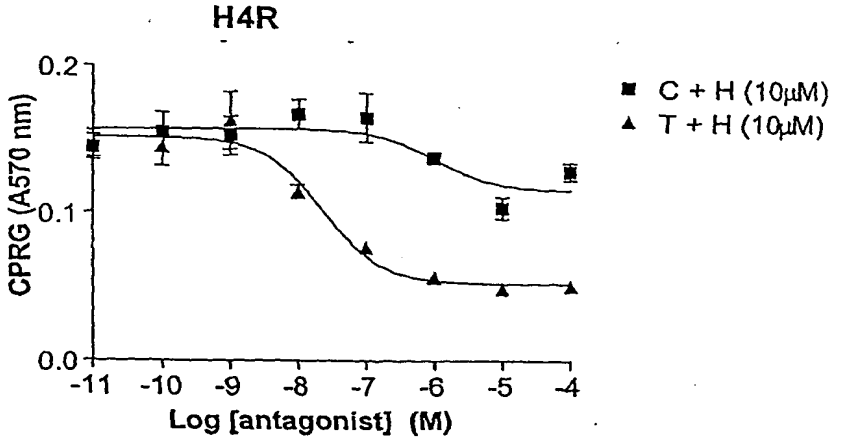

(1) ATGCCAGATACTAATAGCACAATCAATTTATCACTAAGCACTCGTGTTACTTTAGCATTT
M P D T N S T I N L S L S T R V T L A F
TTTATGTCCTTAGTAGCTTTTGCTATAATGCTAGGAAATGCTTTGGTCATTTTAGCTTTT
F M S L V A F A I M L G N A L V I L A F
(121) GTGGTGGACAAAAACCTTAGACATCGAAGTAGTTATTTTTTTCTTAACTTGGCCATCTCT
V V D K N L R H R S S Y F F L N L A I S
GACTTCTTTGTGGGTGTGATCTCCATTCCTTTGTACATCCCTCACACGCTGTTCGAATGG
D F F V G V I S I P L Y I P H T L F E W
241) GATTTTGGAAAGGAAATCTGTGTATTTTGGCTCACTACTGACTATCTGTTATGTACAGCA
D F G K E I C V F W L T T D Y L L C T A
TCTGTATATAACATTGTCCTCATCAGCTATGATCGATACCTGTCAGTCTCAAATGCTGTA
S V Y N I V L I S Y D R Y L S V S N A V
(361) AGTTATAGAACTCAACATACTGGGGTCTTGAAGATTGTTACTCTGATGGTGGCCGTTTGG
S Y R T Q H T G V L K I V T L M V A V W
GTGCTGGCCTTCTTAGTGAATGGGCCAATGATTCTAGTTTCAGAGTCTTGGAAGGATGAA
V L A F L V N G P M I L V S E S W K D E
(481) GGTAGTGAATGTGAACCTGGATTTTTTTCCGAATGGTACATCCTTGCCATCACATCATTC
G S E C E P G F F S E W Y I L A I T S F
TTGGAATTCGTGATCCCAGTCATCTTAGTCGCTTATTTCAACATGAATATTTATTGGAGC
L E F V I P V I L V A Y F N M N I Y W S
(601) CTGTGGAAGCGTGATCATCTCAGTAGGTGCCAAAGCCATCCTGGACTGACTGCTGTCTCT
L W K R D H L S R C Q S H P G L T A V S
TCCAACATCTGTGGACACTCATTCAGAGGTAGACTATCTTCAAGGAGATCTCTTTCTGCA
S N I C G H S F R G R L S S R R S L S A
(721) TCGACAGAAGTTCCTGCATCCTTTCATTCAGAGAGACAGAGGAGAAAGAGTAGTCTCATG
S T E V P A S F H S E R Q R R K S S L M
TTTTCCTCAAGAACCAAGATGAATAGCAATACAATTGCTTCCAAAATGGGTTCCTTCTCC
F S S R T K M N S N T I A S K M G S F S
(841) CAATCAGATTCTGTAGCTCTTCACCAAAGGGAACATGTTGAACTGCTCAGAGCCAGGAGA
Q S D S V A L H Q R E H V E L L R A R R
TTAGCCAAGTCACTGGCCATTCTCTTAGGGGTTTTTGCTGTTTGCTGGGCTCCATATTCT
L A K S L A I L L G V F A V C W A P Y S
(961) CTGTTCACAATTGTCCTTTCATTTTATTCCTCAGCAACAGGTCCTAAATCAGTTTGGTAT
L F T I V L S F Y S S A T G P K S V W Y
AGAATTGCATTTTGGCTTCAGTGGTTCAATTCCTTTGTCAATCCTCTTTTGTATCCATTG
R I A F W L Q W F N S F V N P L L Y P L
(1081) TGTCACAAGCGCTTTCAAAAGGCTTTCTTGAAAATATTTTGTATAAAAAAGCAACCTCTA
C H K R F Q K A F L K I F C I K K Q P L
(1141) CCATCACAACACAGTCGGTCAGTATCTTCTTAA
P S Q H S R S V S S .

Figure 2

A.


H4R Agonists
CPRG (A570 nm)
0.4
0.3
0.2
0.1
0.0
-11 -10 -9 -8 -7 -6 -5 -4
Log [agonist] (M)
Histamine
RαMeHis
Clobenpropit
Thioperamide

B.


H4R
CPRG (A570 nm)
0.2
0.1
0.0
-11 -10 -9 -8 -7 -6 -5 -4
Log [antagonist] (M)
C + H (10μM)
T + H (10μM)

C.

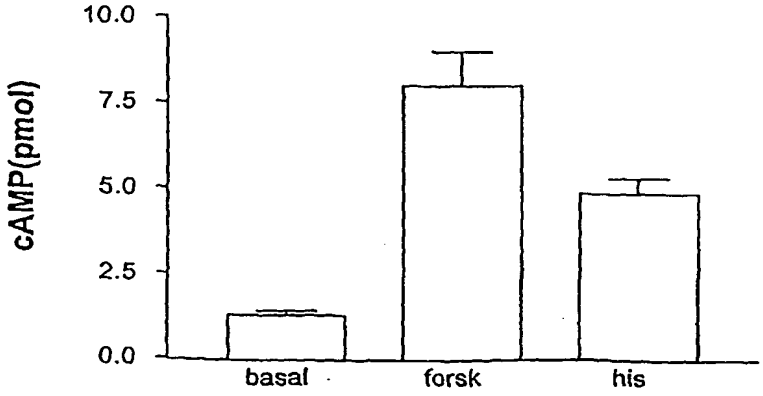

cAMP(pmol)
10.0
7.5
5.0
2.5
0.0
basal
forsk
his

Figure 3

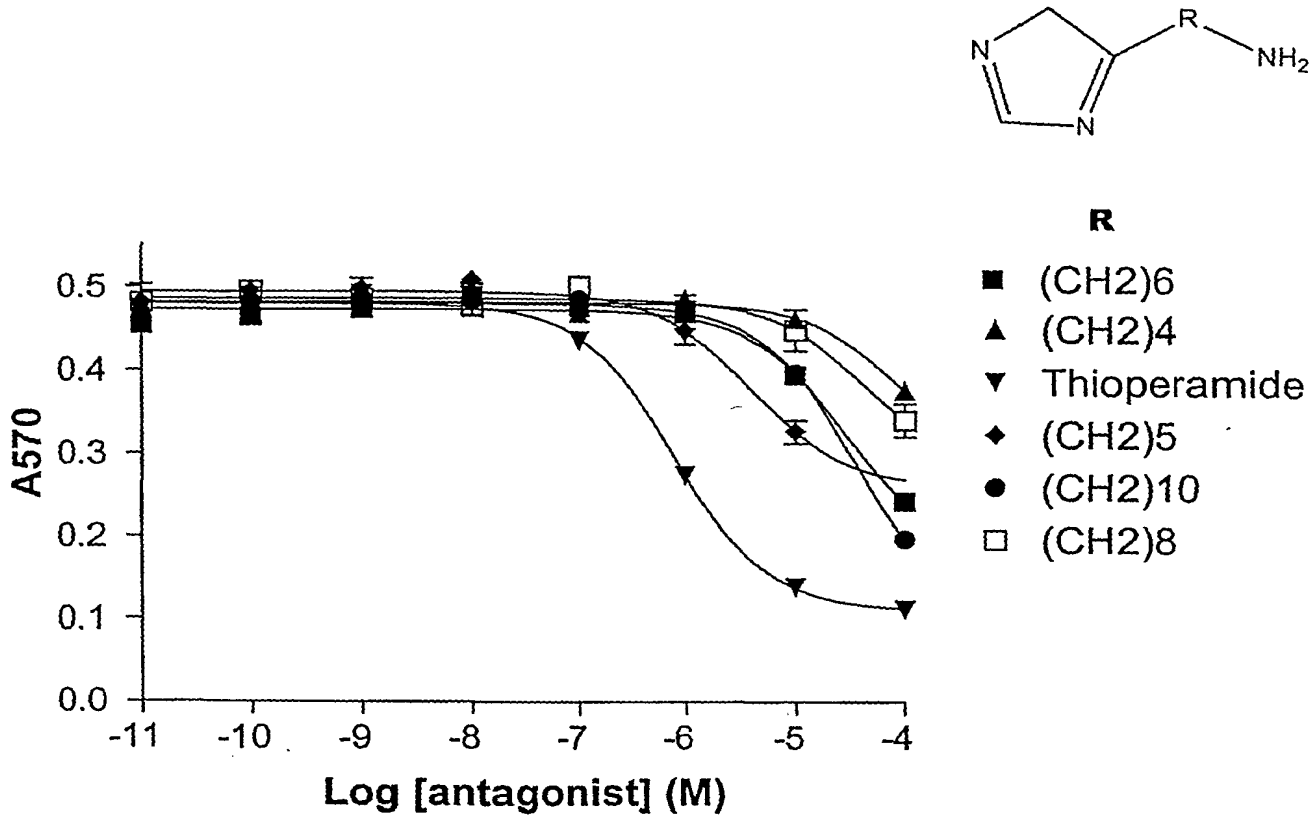
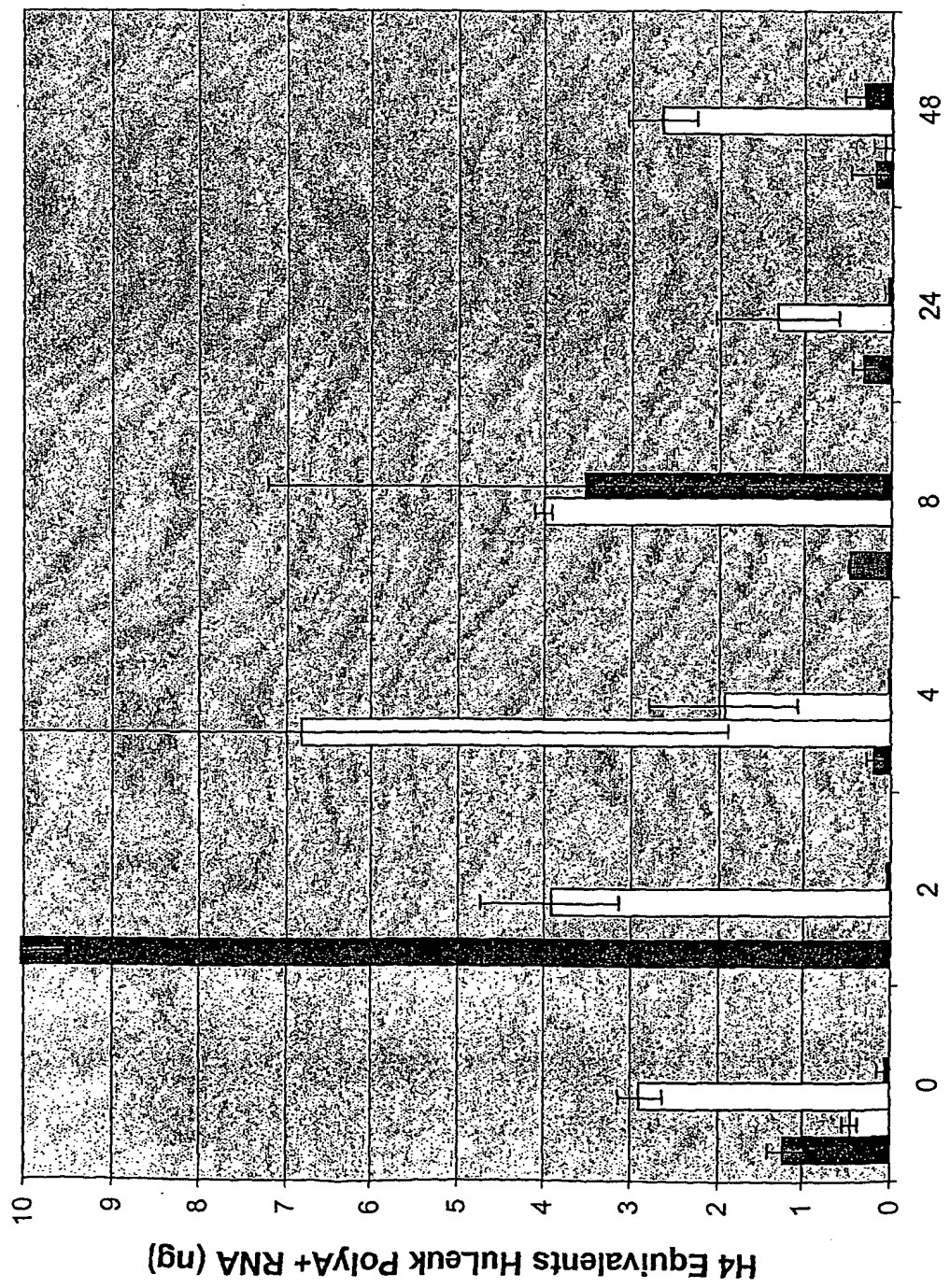

N
R
NH2
N
R
(CH2)6
(CH2)4
Thioperamide
(CH2)5
(CH2)10
(CH2)8
A570
0.5
0.4
0.3
0.2
0.1
0.0
-11
-10
-9
-8
-7
-6
-5
-4
Log [antagonist] (M)

Figure 4



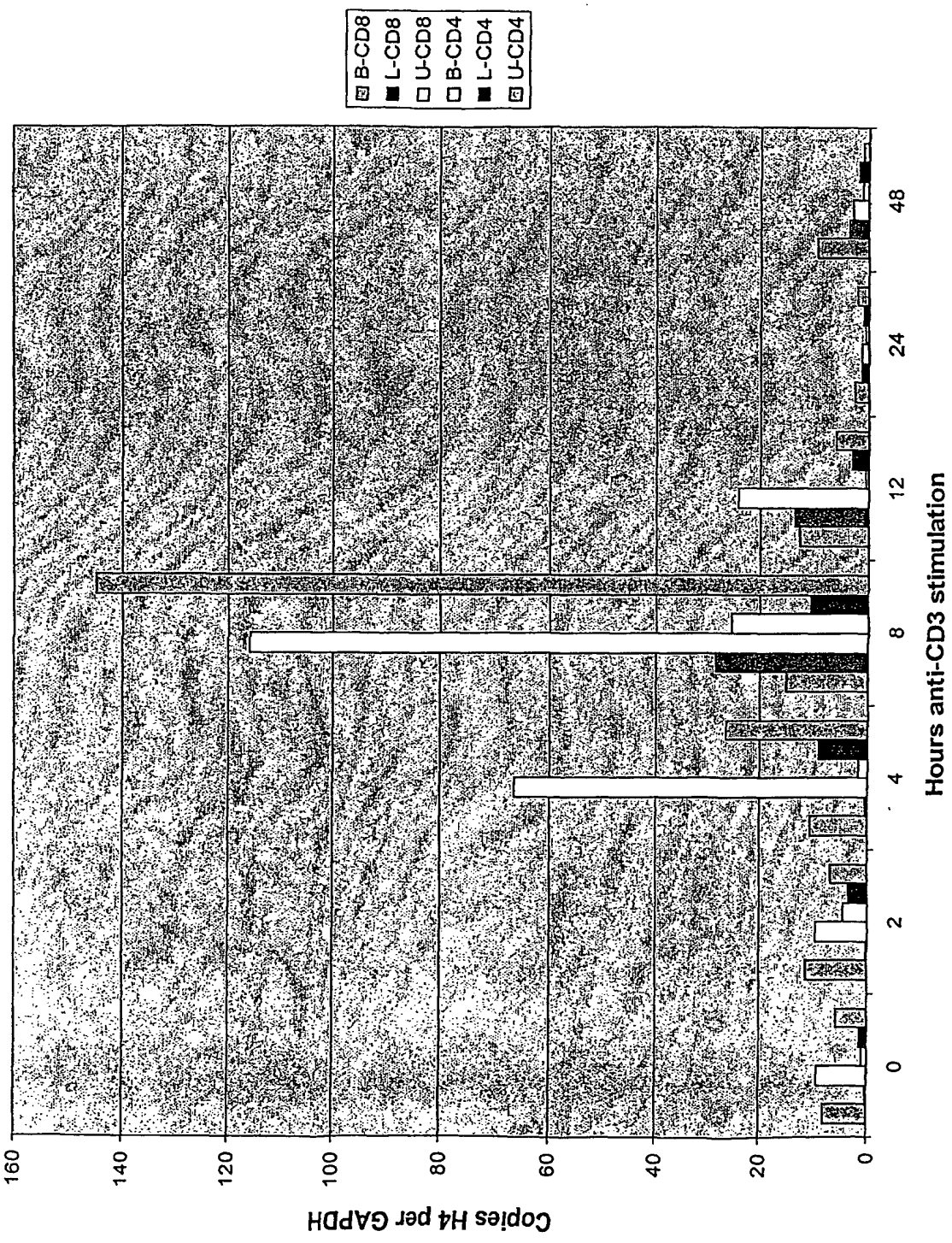


Figure 5

SEQUENCE LISTING

<110> Jones, Philip
Shujian, Wu
Blatcher, Maria
Pausch, Mark

<120> Human Histamine H4 Receptor

<130> 3655/1J100-US3

<140> TBA
<141> Concurrently Herewith

<150> 60/202,151
<151> 2000-05-05

<150> 60/227,567
<151> 2000-08-23

<150> 60/247,855
<151> 2000-11-13

<160> 15

<170> PatentIn version 3.0

<210> 1
<211> 1173
<212> DNA
<213> Homo sapiens

<400> 1
atgccagata ctaatagcac aatcaattta tcactaagca ctcgtgttac tttagcattt 60

tttatgtcct tagtagcttt tgctataatg ctaggaaatg ctttggtcat tttagctttt 120

gtggtggaca aaaaccttag acatcgaagt agttattttt ttcttaactt ggccatctct 180

gacttctttg tgggtgtgat ctccattcct ttgtacatcc ctcacacgct gttcgaatgg 240

gattttggaa aggaaatctg tgtattttgg ctcactactg actatctgtt atgtacagca 300

tctgtatata acattgtcct catcagctat gatcgatacc tgtcagtctc aaatgctgta 360

agttatagaa ctcaacatac tggggtcttg aagattgtta ctctgatggt ggccgtttgg 420

gtgctggcct tcttagtgaa tgggccaatg attctagttt cagagtcttg gaaggatgaa 480

ggtagtgaat gtgaacctgg atttttttcg gaatggtaca tccttgccat cacatcattc 540

ttggaattcg tgatcccagt catcttagtc gcttatttca acatgaatat ttattggagc 600

ctgtggaagc gtgatcatct cagtaggtgc caaagccatc ctggactgac tgctgtctct 660

tccaacatct gtggacactc attcagaggt agactatctt caaggagatc tctttctgca 720

tcgacagaag ttcctgcatc ctttcattca gagagacaga ggagaaagag tagtctcatg 780

ttttcctcaa gaaccaagat gaatagcaat acaattgctt ccaaaatggg ttccttctcc 840

caatcagatt ctgtagctct tcaccaaagg gaacatgttg aactgctcag agccaggaga 900

ttagccaagt cactggccat tctcttaggg gtttttgctg tttgctgggc tccatattct 960

ctgttcacaa ttgtcctttc attttattcc tcagcaacag gtcctaaatc agtttggtat 1020

agaattgcat tttggcttca gtggttcaat tcctttgtca atcctctttt gtatccattg 1080

tgtcacaagc gctttcaaaa ggctttcttg aaaatatttt gtataaaaaa gcaacctcta 1140

ccatcacaac acagtcggtc agtatcttct taa 1173

<210> 2
<211> 390
<212> PRT
<213> Homo sapiens

<400> 2

Met Pro Asp Thr Asn Ser Thr Ile Asn Leu Ser Leu Ser Thr Arg Val
1 5 10 15

Thr Leu Ala Phe Phe Met Ser Leu Val Ala Phe Ala Ile Met Leu Gly
20 25 30

Asn Ala Leu Val Ile Leu Ala Phe Val Val Asp Lys Asn Leu Arg His
35 40 45

Arg Ser Ser Tyr Phe Phe Leu Asn Leu Ala Ile Ser Asp Phe Phe Val
50 55 60

Gly Val Ile Ser Ile Pro Leu Tyr Ile Pro His Thr Leu Phe Glu Trp
65 70 75 80

Asp Phe Gly Lys Glu Ile Cys Val Phe Trp Leu Thr Thr Asp Tyr Leu
85 90 95

Leu Cys Thr Ala Ser Val Tyr Asn Ile Val Leu Ile Ser Tyr Asp Arg
100 105 110

Tyr Leu Ser Val Ser Asn Ala Val Ser Tyr Arg Thr Gln His Thr Gly
115 120 125

Val Leu Lys Ile Val Thr Leu Met Val Ala Val Trp Val Leu Ala Phe
130 135 140

Leu Val Asn Gly Pro Met Ile Leu Val Ser Glu Ser Trp Lys Asp Glu

145 150 155 160

Gly Ser Glu Cys Glu Pro Gly Phe Phe Ser Glu Trp Tyr Ile Leu Ala
165 170 175

Ile Thr Ser Phe Leu Glu Phe Val Ile Pro Val Ile Leu Val Ala Tyr
180 185 190

Phe Asn Met Asn Ile Tyr Trp Ser Leu Trp Lys Arg Asp His Leu Ser
195 200 205

Arg Cys Gln Ser His Pro Gly Leu Thr Ala Val Ser Ser Asn Ile Cys
210 215 220

Gly His Ser Phe Arg Gly Arg Leu Ser Ser Arg Arg Ser Leu Ser Ala
225 230 235 240

Ser Thr Glu Val Pro Ala Ser Phe His Ser Glu Arg Gln Arg Arg Lys
245 250 255

Ser Ser Leu Met Phe Ser Ser Arg Thr Lys Met Asn Ser Asn Thr Ile
260 265 270

Ala Ser Lys Met Gly Ser Phe Ser Gln Ser Asp Ser Val Ala Leu His
275 280 285

Gln Arg Glu His Val Glu Leu Leu Arg Ala Arg Arg Leu Ala Lys Ser
290 295 300

Leu Ala Ile Leu Leu Gly Val Phe Ala Val Cys Trp Ala Pro Tyr Ser
305 310 315 320

Leu Phe Thr Ile Val Leu Ser Phe Tyr Ser Ser Ala Thr Gly Pro Lys
325 330 335

Ser Val Trp Tyr Arg Ile Ala Phe Trp Leu Gln Trp Phe Asn Ser Phe
340 345 350

Val Asn Pro Leu Leu Tyr Pro Leu Cys His Lys Arg Phe Gln Lys Ala
355 360 365

Phe Leu Lys Ile Phe Cys Ile Lys Lys Gln Pro Leu Pro Ser Gln His
370 375 380

Ser Arg Ser Val Ser Ser
385 390

<210> 3
<211> 259
<212> PRT
<213> Artificial

<220>
<223> Consensus sequence

<400> 3

Gly Asn Leu Leu Val Ile Leu Val Ile Leu Arg Thr Lys Lys Leu Arg
1 5 10 15

Thr Pro Thr Asn Ile Phe Ile Leu Asn Leu Ala Val Ala Asp Leu Leu
20 25 30

Phe Leu Leu Thr Leu Pro Pro Trp Ala Leu Tyr Tyr Leu Val Gly Gly
35 40 45

Ser Glu Asp Trp Pro Phe Gly Ser Ala Leu Cys Lys Leu Val Thr Ala
50 55 60

Leu Asp Val Val Asn Met Tyr Ala Ser Ile Leu Leu Leu Thr Ala Ile
65 70 75 80

Ser Ile Asp Arg Tyr Leu Ala Ile Val His Pro Leu Arg Tyr Arg Arg
85 90 95

Arg Arg Thr Ser Pro Arg Arg Ala Lys Val Val Ile Leu Leu Val Trp
100 105 110

Val Leu Ala Leu Leu Leu Ser Leu Pro Pro Leu Leu Phe Ser Trp Val
115 120 125

Lys Thr Val Glu Glu Gly Asn Gly Thr Leu Asn Val Asn Val Thr Val
130 135 140

Cys Leu Ile Asp Phe Pro Glu Glu Ser Thr Ala Ser Val Ser Thr Trp
145 150 155 160

Leu Arg Ser Tyr Val Leu Leu Ser Thr Leu Val Gly Phe Leu Leu Pro
165 170 175

Leu Leu Val Ile Leu Val Cys Tyr Thr Arg Ile Leu Arg Thr Leu Arg
180 185 190

Lys Ala Ala Lys Thr Leu Leu Val Val Val Val Val Phe Val Leu Cys
195 200 205

Trp Leu Pro Tyr Phe Ile Val Leu Leu Leu Asp Thr Leu Cys Leu Ser
210 215 220

Ile Ile Met Ser Ser Thr Cys Glu Leu Glu Arg Val Leu Pro Thr Ala
225 230 235 240

Leu Leu Val Thr Leu Trp Leu Ala Tyr Val Asn Ser Cys Leu Asn Pro
245 250 255

Ile Ile Tyr

<210> 4

<211> 21
<212> DNA
<213> Artificial

<220>
<223> Forward primer

<400> 4
taacttggcc attgacttct t 21

<210> 5
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Reverse Primer

<400> 5
attcgaacag catgtgaggg at 22

<210> 6
<211> 27
<212> DNA
<213> Artificial

<220>
<223> Probe

<220>
<221> modified_base
<222> (1)..(1)
<223> (6-carboxyfluorescein)-t

<220>
<221> modified_base
<222> (27)..(27)
<223> a- (6-carboxy-N,N,N'N'-tetramethylrhodamine)

<400> 6
tacaaaggaa tggagatcac acccaca 27

<210> 7
<211> 29
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 7
aaggatccaa aatgccagat actaatagc 29

<210> 8
<211> 28
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 8
aagtcgactt aagaagatac tgaccgac 28

<210> 9
<211> 43
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 9
aagcttccac catgccagat actaatagca caatcaattt atc 43

<210> 10
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 10
tctagattaa gaagatactg accgactgtg ttg 33

<210> 11
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 11
ggaaggatga aggtagtgaa tg 22

<210> 12
<211> 22

<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 12
cagaatctga ttgggaggaa gg 22

<210> 13
<211> 19
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 13
taacttggcc atctttgac 19

<210> 14
<211> 18
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 14
attcgaacag cgtgtgag 18

<210> 15
<211> 19
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<220>
<221> modified_base
<222> (1)..(1)
<223> 5'-(6-carboxyfluorescein)-t

<400> 15
tacaaaggaa tggagatca 19

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 November 2001 (15.11.2001)

PCT

(10) International Publication Number
WO 01/085786 A3

(51) International Patent Classification[7]: C07K 14/72, C12N 15/16, 5/10, A01K 67/027, C07K 16/28, C12Q 1/68, G01N 33/566

(21) International Application Number: PCT/US01/14527

(22) International Filing Date: 4 May 2001 (04.05.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/202,151	5 May 2000 (05.05.2000)	US
60/227,567	23 August 2000 (23.08.2000)	US
60/247,855	13 November 2000 (13.11.2000)	US

(71) Applicant *(for all designated States except US)*: **WYETH** [US/US]; Five Giralda Farms, Madison, NJ 07940 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **JONES, Philip, G.** [GB/US]; 32 Rocky Brook Road, Cranbury, NJ 08512 (US). **BLATCHER, Maria** [GB/US]; 720 Cox Road, Morristown, NJ 08057 (US). **WU, Shujian** [CN/US]; 972 Ithan Lane, Langhorne, PA 19047 (US). **PAUSCH, Mark, H.** [US/US]; 23 Montgomery Avenue, Rocky Hill, NJ 08553 (US).

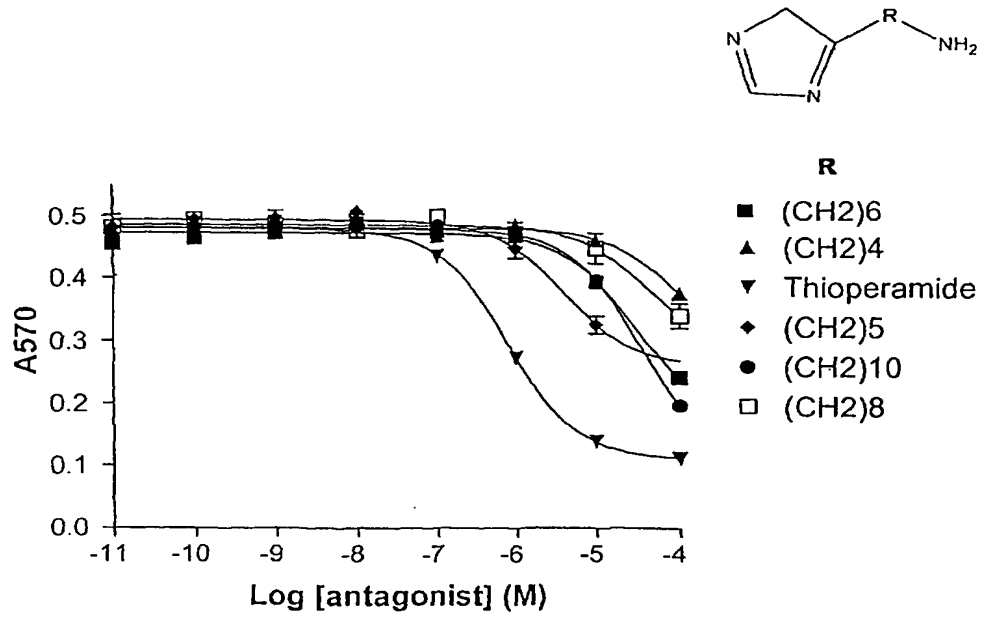
(74) Agent: **EGON, Berg, E.**; Wyeth, Patent Law Department, Five Giralda Farms, Madison, New Jersey 07940 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European

[Continued on next page]

(54) Title: HUMAN HISTAMINE H4 RECEPTOR



(57) Abstract: The present invention discloses the identification of a novel histamine receptor, termed H_4. Amino acid sequences, nucleic acid sequences, vectors, and host cells are also discussed. Additionally, methods of detecting agonists and antagonists for the receptor are disclosed herein.

WO 01/085786 A3

patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
18 April 2002

(48) Date of publication of this corrected version:
12 December 2002

(15) Information about Correction:
see PCT Gazette No. 50/2002 of 12 December 2002, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

HUMAN HISTAMINE H4 RECEPTOR

This application claims priority under 35 U.S.C. § 119 from provisional patent application Serial Nos. 60/202,151, filed May 5, 2000, 60/227,567, filed August 23, 2000, and 60/247,855, filed November 13, 2000; which all are hereby incorporated by reference in their entireties.

FIELD OF THE INVENTION

The present invention discloses the identification of a novel histamine receptor, termed H_4.

BACKGROUND OF THE INVENTION

Histamine produces numerous physiological effects in the body through interaction with one of three different cell surface receptors, classified as H_1, H_2, and H_3. These receptors belong to the guanine nucleotide binding protein coupled receptor (G-protein coupled receptors, GPCR) class.

Stimulation of histamine H_1 receptors produces symptoms that are typically associated with physiological responses to allergic stimuli (Ash and Schild, Br. J. Pharmacol. 1966, 27:427). These effects are blocked by H_1 antagonists such as, for example, diphenhydramine. H_1 antagonists are generally defined as "classical antihistamines". Classical antihistamines are the active ingredient in most over-the-counter allergy medications. Pharmacological studies indicate that agonist activation of these receptors stimulates the inositol phosphate pathway, and thus stimulates formation of inositol triphosphate (IP_3) and diacylglycerol (DAG).

Histamine H_2 receptors have been shown to play a role in gastric acid secretions (Black *et al.*, Nature 1972, 236:385). Histamine H_2 receptor antagonists such as, for example, cimetidine and ranitidine, are often the active ingredient in over-the-counter and prescription drugs that are used to treat duodenal ulcers, gastric ulcers, heartburn, indigestion, and other disorders of the gastrointestinal tract. Activation of histamine H_2 stimulates adenylyl cyclase activity and stimulates formation of cAMP.

Histamine H_3 receptors are a relatively new member of the histamine receptor class. These receptors were originally described as histamine autoreceptors. These receptors were shown to be located on presynaptic histamine nerve terminals and to control the rate of histamine synthesis and release in the brain (Arrang *et al.*, Nature 1983, 302:832). Recent studies indicate that H_3 receptors also are present on non-histamine presynaptic nerve terminals. H_3 receptors may modulate the synthesis and release of other neurotransmitters such as dopamine, serotonin, acetylcholine, and norepinephrine. H_3 receptors also have been found in peripheral tissues. Histamine activation of H_3 receptor inhibits adenylyl cyclase activity and formation of cAMP.

Multiple pharmacological studies have indicated the presence of H_3 receptor subtypes (Leurs *et al.* JPET 1996, 276:1009-1015; Cumming and Gjede, Brain Res. 1994, 641:203-207; Calpham and Kilpatrick, Br. J. Pharmacol 1996, 107:919-923; Schworer *et al.* Naunyn-Schmiedeberg's Arch. Pharmacol. 1994, 350:375-379; Schkicker *et al.*, Naunyn-Schmiedeberg's Arch. Pharmacol. 1996, 353:482-488). Additionally, pharmacological characterization of a histamine receptor on eosinophils describes the greater potency of histamine compared to R-α-methylhistamine (Raible, *et al.* Am. J. Respir. Crit. Care.Med 1994 , 149:1506-1511.). However, the existence of these receptor subtypes has yet to be substantiated by molecular biological techniques.

SUMMARY OF THE INVENTION

The present invention contemplates an isolated histamine H_4 receptor protein having an amino acid sequence at least 51% identical or comprising at least 10 contiguous amino acids from the sequence depicted in SEQ ID NO:2. The H_4 receptor protein binds ligands comprising an imidazole and an amine, which imidazole and amine are attached by an alkyl chain, where the rank order of efficacy of modulation of second messenger formation of the ligands at the H_4 receptor protein is 5>6=10>8=4, where the number represents the number of carbons in the alkyl chain. In one embodiment, upon binding histamine or a histamine agonist the receptor protein inhibits second messenger formation. Preferably the second messenger is cAMP.

The present application also discloses an isolated H_4 receptor protein having an amino acid sequence with at least 95% sequence identity to human H_4 receptor protein having

an amino acid sequence as depicted in SEQ ID NO: 2. In a specific embodiment, the protein is a human H_4 receptor protein, *e.g.*, having an amino acid sequence as depicted in SEQ ID NO:2 or an allelic variant of that sequence.

Also provided are nucleic acids, *e.g.*, cDNAs encoding the H_4 receptor proteins as discussed above. In another embodiment, an isolated nucleic acid encoding an H_4 receptor protein, which nucleic acid hybridizes under stringent conditions to a nucleic acid having a sequence of at least 20 nucleotides identical to a corresponding nucleotide sequence of the same number of bases in SEQ ID NO:1 or its complement.

The isolated nucleic acids encoding H_4 receptors can be part of vectors, *e.g.*, for cloning, expression, and/or expansion. An expression vector comprises the nucleic acid encoding the H_4 receptor protein operably associated with an expression control sequence. The invention further provides host cells and non-human transgenic animals containing such an expressible vector, and methods for producing an H_4 receptor polypeptide using such host cells.

In addition, the invention provides an isolated nucleic acid, such as a primer or probe, of at least 10 bases having a nucleotide sequence identical to a corresponding nucleotide sequence of the same number of bases in SEQ ID NO:1, or its complement. The invention also provides an antibody that specifically binds an H_4 receptor.

The present invention further contemplates a method for detecting expression of H_4 receptor, which method comprises detecting mRNA encoding H_4 receptor in a sample from a cell suspected of expressing H_4 receptor or detecting the H_4 receptor protein with an antibody of the invention.

The present invention also contemplates an assay system for identifying H_4 receptor ligands. The assay system comprises a sufficient number of transformed host cells to be able to detect an alteration in second messenger accumulation. Preferably, the second messenger is cAMP.

The present invention also contemplates a method for identifying a test compound that antagonizes or agonizes histamine H_4 receptors. The method comprises detecting an alteration in the level of a second messenger in the assay system contacted with the test compound. In the method, an increase in the level of the second messenger indicates that the test compound antagonizes the H_4 receptor. A decrease in the level of the second messenger indicates that the test compound agonizes the H_4 receptor.

The present invention also discloses an isolated nucleic acid that specifically hybridizes under highly stringent conditions to the complement of the sequence depicted in SEQ ID NO:1, where the nucleic acid encodes a histamine H_4 receptor protein.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1. Sequence of the human H_4 cDNA (SEQ ID NO: 1) and the deduced amino acid sequence of the protein it encodes (SEQ ID NO:2). The TM domains are denoted by lines.

Figures 2A, 2B, and 2C. (A) Agonist assay, Square: histamine; triangle: R-α-methylhistamine; inverted triangle: clobenpropit; diamond: thioperamide. **(B)** Antagonist assay. Assay were conducted in duplicate and presented as the average ± SEM. **(C)** Activity in mammalian cells.

Figure 3. Antagonist activity of a series of histamine analogs. Square: clobenpropit; triangle: thioperamide. Assay were conducted in duplicate and presented as the average ± SEM.

Figure 4. H_4 expression in from 6 human T cell clones derived from a single human donor. Three of these clones were CD4+ cell clones (RG4.3B, RG4.3A, and RG4.3), two were CD8+ cell clones (RG8.1C and RG8.1A), and one was an NKT cell clone (RG1).

Figure 5. H_4 expression in bulk populations of CD4+ and CD8+ T cells isolated from three different individuals. B-CD8 and B-CD4 were CD8+ and CD4+ cells, respectively, obtained from individual 1. L-CD8 and L-CD4 were CD8+ and CD4+ cells, respectively, obtained from individual 2. U-CD8 and U-CD4 were CD8+ and CD4+ cells, respectively, obtained from individual 3.

DETAILED DESCRIPTION

The present invention is based, in part, on discovery of a novel histamine receptor, which has been termed H_4. The new histamine receptor was cloned from a human heart library using H_3 specific primers, but proved to be a new histamine receptor. Yeast and human endothelial kidney (HEK) cells were transformed with a human H_4 expression vector. Transformed cells in multi-well plates were treated with test compounds, and regulation of the intracellular second messenger cyclic adenosine monophosphate (cAMP) formation was

determined. Modulation of cAMP formation is ligand and concentration dependent.

The nucleic acid and protein sequences of H_4 shows homology to known G-protein coupled receptors. Specifically, the H_4 protein shows homology to biogenic-amine G-protein coupled receptors. As in other biogenic amine receptors, the present protein contains conserved aspartate residues within transmembrane (TM) domains 2 and 3 (at positions 61 and 94). A DY motif (D94 and Y95) in TM3 is found only in histamine and muscarinic receptors. A DRY sequence motif that exists at the cytoplasmic interface of TM3 of all GPCRs also can be found in the present sequence (amino acids 111- 113). Also present are two conserved cysteine residues in the first and second extracellular loops (position 88 and 164), which are predicted to form a disulfide bond; two conserved tryptophan residues in TM4 and TM6 (positions 140 and 316); and conserved proline residues in TM5, TM6, and TM7 (positions 186, 318, and 355, respectively). The putative TM domains are shown in Figure 1. H_4 shares sequence identity with human H_3 receptors. Sequence comparison indicates an overall sequence identity of about 44% and overall sequence similarity of about 51%.

The present invention also contemplates an assay method and system for identifying selective H_4 receptor ligands. The method involves detecting binding of a test compound to isolated cell membranes containing the histamine H_4 receptor. The assay system comprises transformed host cells that express H_4 receptors, where the number of cells in the assay system is sufficient to detect an alteration in second messenger accumulation. The test system also includes an appropriate cell culture medium to permit cell growth and viability, and preferably tissue culture plates or arrays containing the host cells in cell culture medium. In a specific embodiment, the second messenger that is detected is cAMP. In a further embodiment, the receptor is a human receptor.

The invention also discloses a method for identifying a test compound that antagonizes or agonizes histamine H_4 receptors. The method comprises detecting an increase (antagonist) or decrease (agonist) in the level of a second messenger in the assay system when contacted with the test compound.

Thus, the present invention advantageously provides H_4 protein, including fragments, derivatives, and analogs of H_4; H_4 nucleic acids, including oligonucleotide primers and probes, and H_4 regulatory sequences (especially an H_4 primer and splice sites with introns); H_4-specific antibodies; and related methods of using these materials to detect the presence of H_4

proteins or nucleic acids, H_4 binding partners, and in screens for agonists and antagonists of H_4.

The following sections of the application, which are delineated by headings (in bold) and sub-headings (in bold italics), which cover these three aspects of the invention, are provided for clarity, and not by way of limitation.

General Definitions

As used herein, the term "isolated" means that the referenced material is removed from the environment in which it is normally found. Thus, an isolated biological material can be free of cellular components, *i.e.*, components of the cells in which the material is found or produced in nature. In the case of nucleic acid molecules, an isolated nucleic acid includes a PCR product, an isolated mRNA, a cDNA, or a restriction fragment. In another embodiment, an isolated nucleic acid is preferably excised from the chromosome in which it may be found, and more preferably is no longer joined to non-regulatory, non-coding regions, or to other genes, located upstream or downstream of the gene contained by the isolated nucleic acid molecule when found in the chromosome. In yet another embodiment, the isolated nucleic acid lacks one or more introns. Isolated nucleic acid molecules include sequences inserted into plasmids, cosmids, artificial chromosomes, and the like. Thus, in a specific embodiment, a recombinant nucleic acid is an isolated nucleic acid. An isolated protein may be associated with other proteins or nucleic acids, or both, with which it associates in the cell, or with cellular membranes if it is a membrane-associated protein. A protein expressed from a vector in a cell, particularly a cell in which the protein is normally not expressed is also a regarded as isolated. An isolated organelle, cell, or tissue is removed from the anatomical site in which it is found in a cell or an organism. An isolated material may be, but need not be, purified.

The term "purified" as used herein refers to material that has been isolated under conditions that reduce or eliminate the presence of unrelated materials, *i.e.*, contaminants, including native materials from which the material is obtained. For example, a purified protein is preferably substantially free of other proteins or nucleic acids with which it is associated in a cell; a purified nucleic acid molecule is preferably substantially free of proteins or other unrelated nucleic acid molecules with which it can be found within a cell. As used herein, the term "substantially free" is used operationally, in the context of analytical testing of the material. Preferably, purified material substantially free of contaminants is at least 50% pure; more

preferably, at least 90% pure; and more preferably still at least 99% pure. Purity can be evaluated by chromatography, gel electrophoresis, immunoassay, composition analysis, biological assay, and other methods known in the art.

Methods for purification are well-known in the art. For example, nucleic acids can be purified by precipitation, chromatography (including preparative solid phase chromatography, oligonucleotide hybridization, and triple helix chromatography), ultracentrifugation, and other means. Polypeptides and proteins can be purified by various methods including, without limitation, preparative disc-gel electrophoresis, isoelectric focusing, HPLC, reversed-phase HPLC, gel filtration, ion exchange and partition chromatography, precipitation and salting-out chromatography, extraction, and countercurrent distribution. For some purposes, it is preferable to produce the polypeptide in a recombinant system in which the protein contains an additional sequence tag that facilitates purification, such as, but not limited to, a polyhistidine sequence, or a sequence that specifically binds to an antibody, such as FLAG and GST. The polypeptide can then be purified from a crude lysate of the host cell by chromatography on an appropriate solid-phase matrix. Alternatively, antibodies produced against the protein or against peptides derived therefrom can be used as purification reagents. Cells can be purified by various techniques, including centrifugation, matrix separation (*e.g.*, nylon wool separation), panning and other immunoselection techniques, depletion (*e.g.*, complement depletion of contaminating cells), and cell sorting (*e.g.*, fluorescence activated cell sorting [FACS]). Other purification methods are possible. A purified material may contain less than about 50%, preferably less than about 75%, and most preferably less than about 90%, of the cellular components with which it was originally associated. The "substantially pure" indicates the highest degree of purity which can be achieved using conventional purification techniques known in the art.

In a specific embodiment, the term "about" or "approximately" means within a scientifically acceptable error range for a given value relative to the precision with which the value is or can be measured, *e.g.*, within 20%, preferably within 10%, and more preferably within 5% of a given value or range. Alternatively, particularly with biological systems, the term can mean within an order of magnitude, preferably within 5-fold and more preferably within 2-fold of a given value.

A "sample" as used herein refers to a biological material which can be tested for

the presence of H_4 protein or H_4 nucleic acids. Such samples can be obtained from animal subjects, such as humans and non-human animals, and include tissue, especially muscle, biopsies, blood and blood products; plural effusions; cerebrospinal fluid (CSF); ascites fluid; and cell culture.

Non-human animals include, without limitation, laboratory animals such as mice, rats, rabbits, hamsters, guinea pigs, etc.; domestic animals such as dogs and cats; and, farm animals such as sheep, goats, pigs, horses, and cows.

The use of italics indicates a nucleic acid molecule; normal text indicates the polypeptide or protein.

The term "ligand" refers to a compound that recognizes and binds to a receptor binding site. In a specific embodiment, the ligand binds to the histamine H_4 receptors of the invention. Upon binding to the receptor, the ligand may produce agonist or antagonist functional effects.

The term "agonist" refers to a ligand that binds to the receptor and produces a functional effect similar to that produced by the endogenous ligand for the receptor. In a specific embodiment, the agonist at the histamine H_4 receptor produces an effect similar to that produced by histamine, the endogenous ligand (histamine) for the H_4 receptor. Examples of such agonists include, but are not limited to, R-α-methyl histamine and imetit.

The term "antagonist" refers to a ligand that binds to the receptor and blocks a functional effect produced by an agonist for the receptor or the endogenous ligand of the receptor. Examples of such antagonists include, but are not limited to, thioperamide.

The term "selective" refers to the ability of a histamine H_4 agonist or antagonist to elicit a response from the H_4 receptor while eliciting minimal responses from another receptor. Stated differently, a selective H_4 agonist may be a potent agonist for the H_4 receptor while agonizing another receptor, such as another G-protein coupled receptor and particularly another histamine receptor, poorly or not at all.

The term "ability to elicit a response" refers to the ability of a H_4 agonist or antagonist ligand to agonize or antagonize H_4 receptor activity.

As used herein the term "transformed cell" refers to a modified host cell that expresses a functional H_4 receptor expressed from a vector encoding the histamine receptor. Any cell can be used, preferably a mammalian cell, and more preferably a HEK cell.

A "functional histamine receptor" is a receptor that binds histamine or H_4 agonists and transduces a signal upon such binding. Preferably the H_4 receptor is a human H_4. Preferably, the signal that is transduced is accumulation of a second messenger, preferably cAMP. Histamine H_4 receptors may be derived from a variety of sources, including mammal, *e.g.*, human, bovine, porcine, and canine; and avian.

The cells of the invention are particularly suitable for an assay system for histamine H_4 receptor ligands that modulate second messenger accumulation. An "assay system" is one or more collections of such cells, *e.g.*, in a microwell plate or some other culture system. To permit evaluation of the effects of a test compound on the cells, the number of cells in a single assay system is sufficient to express a detectable amount of the regulated second messenger accumulation at least under conditions of maximum second messenger accumulation.

A "second messenger" is an intracellular molecule or ion, where formation and/or accumulation of the second messenger is regulated by activation of cellular membranes. In one embodiment, cellular membranes contain G-protein coupled receptor, ion channels, and tyroinse kinase receptors. In the context of this invention, the cellular membrane is a G-protein coupled receptor, preferably a histamine H_4 receptor. In a specific embodiment, the second messenger is one or more of cAMP, cGMP, inositol phosphate, DAG, and ions such as calcium and potassium. Preferably, the second messenger is cAMP.

A "test compound" or "candidate compound" is any molecule that can be tested for its ability to bind H_4 receptors, and preferably modulate second messenger accumulation through the H_4 receptor, as set forth herein. A compound that binds, and preferably modulates H_4 is a "lead compound" suitable for further testing and development as an H_4 agonist or antagonist.

As used herein, the term "provide" refers to supplying the compounds or pharmaceutical compositions of the present invention to cells or to an animal, preferably a human, in any form. For example, a prodrug form of the compounds may be provided the subject, which then is metabolized to the compound in the body.

H_4 Receptor

H_4 receptor, as defined herein, refers to a polypeptide having about 390 amino acids.

The protein is transcribed from a nucleic acid sequence that is about 1173 base pairs in length. The H_4 protein has significant homology to the H_3 receptor. Thus, H_4 refers to orthologs and allelic variants, *e.g.*, a protein having greater than about 50%, preferably greater than 80%, more preferably still greater than 90%, and even more preferably greater than 95% overall sequence identity to SEQ ID NO: 2. Allelic variants may differ from 1 to about 5 amino residues from SEQ ID NO:2. In a specific embodiment, H_4 has an amino acid sequence as shown in SEQ ID NO: 2.

Sequence comparison studies between human H_4 protein and human H_3 protein, indicates sequence identity of about 44%. The predicted protein sequence contains residues which are characteristic of the class of biogenic amine receptors. As in other biogenic amine receptors, the present protein contains conserved aspartate residues within TM2 and TM3 (at positions 61 and 94). A DY motif, found only in histamine and muscarinic receptors, is found at positions 94-95. A DRY sequence motif that exists at the cytoplasmic interface of TM3 of all GPCRs also can be found in amino acids 111-113. Also present are two conserved cysteine residues in the first and second extracellular loops (position 88 and 164), two conserved tryptophan residues in TM4 and TM6 (positions 140 and 316); and conserved proline residues in TM5, TM6, and TM7 (positions 186, 318, and 355, respectively).

H_4 receptors, like H_2 and H_3 receptors, modulates adenylyl cyclase activity. Therefore, the receptor modulates accumulation of the intracellular messenger cAMP. Modulation of H_4 receptors may be a treatment for transplanted organ rejection, asthma, allergies, and autoimmune pathologies, such as, but not limited to, multiple sclerosis, type I diabetes, and rheumatoid arthritis. Additionally, CNS functions such as cognitive and memory defects may be treated.

H_4 fragments, derivatives, and analogs can be characterized by one or more of the characteristics of H_4 protein. In a specific embodiment, in order to develop the specific C-terminal and N-terminal H_4 antibodies, antibodies can be raised against extracellular or cytoplasmic portions of the H_4 protein, preferably or antigenic peptides identified using a hydrophobicity profile or other algorithms.

Analogs and derivatives of the H_4 receptor of the invention have the same or homologous characteristics of H_4 as set forth above. For example, a truncated form of H_4 can be provided. Such a truncated form includes H_4 with a either an N-terminal, C-terminal, or internal deletion. In a specific embodiment, the derivative is functionally active, *i.e.*, capable of exhibiting one or more functional activities associated with a full-length, wild-type H_4 of the invention. Such functions include, but are not limited to, inhibition of adenylyl cyclase activity and cAMP formation. Alternatively, a H_4 chimeric fusion protein can be prepared in which the H_4 portion of the fusion protein has one or more characteristics of H_4. Such fusion proteins include fusions of the H_4 receptor with a marker polypeptide, such as FLAG, a histidine tag, a myc tag, or glutathione-S-transferase (GST). Alternatively, the H_4 receptor can be fused with an expression-related peptide, such as yeast α-mating factor, a heterogeneous signal peptide, or a peptide that renders the protein more stable upon expression. H_4 can also be fused with a unique phosphorylation site for labeling.

H_4 analogs can be made by altering encoding nucleic acid sequences by substitutions, additions or deletions that provide for functionally similar molecules, *i.e.*, molecules that perform one or more H_4 functions. In a specific embodiment, an analog of H_4 is a sequence-conservative variant of H_4. In another embodiment, an analog of H_4 is a function-conservative variant. In yet another embodiment, an analog of H_4 is an allelic variant or a homologous variant from another species. In a specific embodiment, human variants of H_4 are described.

H_4 derivatives include, but are by no means limited to, phosphorylated H_4, glycosylated H_4, methylated H_4, and other H_4 proteins that are otherwise chemically modified. H_4 derivatives also include labeled variants, *e.g.*, radio-labeled with iodine (or, as pointed out above, phosphorous); a detectable molecule, such as but by no means limited to biotin, a chelating group complexed with a metal ion, a chromophore or fluorophore, a gold colloid, or a particle such as a latex bead; or attached to a water soluble polymer.

Cloning and Expression of H_4

The present invention contemplates analysis and isolation of a gene encoding a functional or mutant H_4, including a full length, or naturally occurring form of H_4, and any antigenic fragments thereof from any source, preferably human. It further contemplates

expression of functional or mutant H_4 protein for evaluation, diagnosis, or therapy.

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, *e.g.*, Sambrook, Fritsch & Maniatis, Molecular Cloning: A Laboratory Manual, Second Edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (herein "Sambrook *et al.*, 1989"); DNA Cloning: A Practical Approach, Volumes I and II (D.N. Glover ed. 1985); Oligonucleotide Synthesis (M.J. Gait ed. 1984); Nucleic Acid Hybridization [B.D. Hames & S.J. Higgins eds. (1985)]; Transcription And Translation [B.D. Hames & S.J. Higgins, eds. (1984)]; Animal Cell Culture [R.I. Freshney, ed. (1986)]; Immobilized Cells And Enzymes [IRL Press, (1986)]; B.Perbal, A Practical Guide To Molecular Cloning (1984); F.M. Ausubel *et al.* (eds.), Current Protocols in Molecular Biology, John Wiley & Sons, Inc. (1994).

Molecular Biology - Definitions

"Amplification" of DNA as used herein denotes the use of polymerase chain reaction (PCR) to increase the concentration of a particular DNA sequence within a mixture of DNA sequences. For a description of PCR see Saiki *et al.*, Science, 239:487, 1988.

As used herein, "sequence-specific oligonucleotides" refers to related sets of oligonucleotides that can be used to detect allelic variations or mutations in the H_4 gene.

The nucleic acid molecules (polynucleotides) herein may be flanked by natural regulatory (expression control) sequences, or may be associated with heterologous sequences, including promoters, internal ribosome entry sites (IRES) and other ribosome binding site sequences, enhancers, response elements, suppressors, signal sequences, polyadenylation sequences, introns, 5'- and 3'- non-coding regions, and the like. The nucleic acids may also be modified by many means known in the art. Non-limiting examples of such modifications include methylation, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, and internucleotide modifications such as, for example, those with uncharged linkages (*e.g.*, methyl phosphonates, phosphotriesters, phosphoroamidates, carbamates, etc.) and with charged linkages (*e.g.*, phosphorothioates, phosphorodithioates, etc.). Polynucleotides may contain one or more additional covalently linked moieties, such as, for example, proteins (*e.g.*, nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.), intercalators (*e.g.*, acridine, psoralen, etc.), chelators (*e.g.*, metals, radioactive metals, iron, oxidative metals, etc.), and

alkylators. The polynucleotides may be derivatized by formation of a methyl or ethyl phosphotriester or an alkyl phosphoramidate linkage. Furthermore, the polynucleotides herein may also be modified with a label capable of providing a detectable signal, either directly or indirectly. Exemplary labels include radioisotopes, fluorescent molecules, biotin, and the like.

A "coding sequence" or a sequence "encoding" an expression product, such as a RNA, polypeptide, protein, or enzyme, is a nucleotide sequence that, when expressed, results in the production of that RNA, polypeptide, protein, or enzyme, *i.e.*, the nucleotide sequence encodes an amino acid sequence for that polypeptide, protein or enzyme. A coding sequence for a protein may include a start codon (usually ATG) and a stop codon.

The term "gene", also called a "structural gene" means a DNA sequence that codes for or corresponds to a particular sequence of amino acids which comprise all or part of one or more proteins or enzymes, and may or may not include introns and regulatory DNA sequences, such as promoter sequences, 5'-untranslated region, or 3'-untranslated region which affect for example the conditions under which the gene is expressed.

A "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site (conveniently defined for example, by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase. The present invention includes the H_4 receptor gene promoter found in the genome, which can be operatively associated with a H_4 coding sequence with a heterologous coding sequence.

The term "host cell" means any cell of any organism that is selected, modified, transformed, grown, or used or manipulated in any way, for the production of a substance by the cell, for example the expression by the cell of a gene, a DNA or RNA sequence, a protein or an enzyme. Host cells can further be used for screening or other assays, as described *infra*.

A coding sequence is "under the control of" or "operatively associated with" transcriptional and translational control sequences in a cell when RNA polymerase transcribes the coding sequence into mRNA, which is then trans-RNA spliced (if it contains introns) and

translated, in the case of mRNA, into the protein encoded by the coding sequence.

The terms "express" and "expression" mean allowing or causing the information in a gene or DNA sequence to become manifest, for example producing a protein by activating the cellular functions involved in transcription and translation of a corresponding gene or DNA sequence. A DNA sequence is expressed in or by a cell to form an "expression product" such as a protein. The expression product itself , *e.g.,* the resulting protein, may also be said to be "expressed" by the cell. An expression product can be characterized as intracellular, extracellular, transmembrane, or secreted. The term "intracellular" means something that is inside a cell. The term "extracellular" means something that is outside a cell. The H_4 receptor is a seven transmembrane protein with intracellular and extracellular domains. A substance is "secreted" by a cell if it appears in significant measure outside the cell, from somewhere on or inside the cell.

The term "transfection" means the introduction of a foreign nucleic acid into a cell. The term "transformation" means the introduction of a "foreign" (*i.e.,* extrinsic or extracellular) gene, DNA or RNA sequence to a host cell, so that the host cell will express the introduced gene or sequence to produce a desired substance, typically a protein or enzyme coded by the introduced gene or sequence. The introduced gene or sequence may also be called a "cloned" or "foreign" gene or sequence, may include regulatory or control sequences, such as start, stop, promoter, signal, secretion, or other sequences used by a cell's genetic machinery. The gene or sequence may include nonfunctional sequences or sequences with no known function. A host cell that receives and expresses introduced DNA or RNA has been "transformed" and is a "transformant" or a "clone." The DNA or RNA introduced to a host cell can come from any source, including cells of the same genus or species as the host cell, or cells of a different genus or species.

The terms "vector", "cloning vector" and "expression vector" mean the vehicle by which a DNA or RNA sequence (*e.g.,* a foreign gene) can be introduced into a host cell, so as to transform the host and promote expression (*e.g.,* transcription and translation) of the introduced sequence. Vectors include plasmids, phages, viruses, etc.; they are discussed in greater detail below.

Vectors typically comprise the DNA of a transmissible agent, into which foreign DNA is inserted. A common way to insert one segment of DNA into another segment of DNA

involves the use of enzymes called restriction enzymes that cleave DNA at specific sites (specific groups of nucleotides) called restriction sites. A "cassette" refers to a DNA coding sequence or segment of DNA that codes for an expression product that can be inserted into a vector at defined restriction sites. The cassette restriction sites are designed to ensure insertion of the cassette in the proper reading frame. Generally, foreign DNA is inserted at one or more restriction sites of the vector DNA, and then is carried by the vector into a host cell along with the transmissible vector DNA. A segment or sequence of DNA having inserted or added DNA, such as an expression vector, can also be called a "DNA construct." A common type of vector is a "plasmid", which generally is a self-contained molecule of double-stranded DNA, usually of bacterial origin, that can readily accept additional (foreign) DNA and which can readily introduced into a suitable host cell. A plasmid vector often contains coding DNA and promoter DNA and has one or more restriction sites suitable for inserting foreign DNA. Coding DNA is a DNA sequence that encodes a particular amino acid sequence for a particular protein or enzyme. Promoter DNA is a DNA sequence which initiates, regulates, or otherwise mediates or controls the expression of the coding DNA. Promoter DNA and coding DNA may be from the same gene or from different genes, and may be from the same or different organisms. A large number of vectors, including plasmid and fungal vectors, have been described for replication and/or expression in a variety of eukaryotic and prokaryotic hosts. Non-limiting examples include pKK plasmids (Clonetech), pUC plasmids, pET plasmids (Novagen, Inc., Madison, WI), pRSET or pREP plasmids (Invitrogen, San Diego, CA), or pMAL plasmids (New England Biolabs, Beverly, MA), and many appropriate host cells, using methods disclosed or cited herein or otherwise known to those skilled in the relevant art. Recombinant cloning vectors will often include one or more replication systems for cloning or expression, one or more markers for selection in the host, *e.g.*, antibiotic resistance, and one or more expression cassettes.

The term "expression system" means a host cell and compatible vector under suitable conditions, *e.g.*, for the expression of a protein coded for by foreign DNA carried by the vector and introduced to the host cell. Common expression systems include E. coli host cells and plasmid vectors, insect host cells and Baculovirus vectors, and mammalian host cells and vectors.

The term "heterologous" refers to a combination of elements not naturally occurring. For example, heterologous DNA refers to DNA not naturally located in the cell, or

in a chromosomal site of the cell. Preferably, the heterologous DNA includes a gene foreign to the cell. A heterologous expression regulatory element is such an element operatively associated with a different gene than the one it is operatively associated with in nature. In the context of the present invention, an H_4 gene is heterologous to the vector DNA in which it is inserted for cloning or expression, and it is heterologous to a host cell containing such a vector, in which it is expressed, *e.g.*, a HEK cell.

The terms "mutant" and "mutation" mean any detectable change in genetic material, *e.g.*, DNA, or any process, mechanism, or result of such a change. This includes gene mutations, in which the structure (*e.g.*, DNA sequence) of a gene is altered, any gene or DNA arising from any mutation process, and any expression product (*e.g.*, protein or enzyme) expressed by a modified gene or DNA sequence. The term "variant" may also be used to indicate a modified or altered gene, DNA sequence, enzyme, cell, etc., *i.e.*, any kind of mutant.

"Sequence-conservative variants" of a polynucleotide sequence are those in which a change of one or more nucleotides in a given codon position results in no alteration in the amino acid encoded at that position.

"Function-conservative variants" are those in which a given amino acid residue in a protein or enzyme has been changed without altering the overall conformation and function of the polypeptide, including, but not limited to, replacement of an amino acid with one having similar properties (such as, for example, polarity, hydrogen bonding potential, acidic, basic, hydrophobic, aromatic, and the like). Amino acids with similar properties are well known in the art. For example, arginine, histidine and lysine are hydrophilic-basic amino acids and may be interchangeable. Similarly, isoleucine, a hydrophobic amino acid, may be replaced with leucine, methionine or valine. Such changes are expected to have little or no effect on the apparent molecular weight or isoelectric point of the protein or polypeptide. Amino acids other than those indicated as conserved may differ in a protein or enzyme so that the percent protein or amino acid sequence similarity between any two proteins of similar function may vary and may be, for example, from 70% to 99% as determined according to an alignment scheme such as by the Cluster Method, wherein similarity is based on the MEGALIGN algorithm. A "function-conservative variant" also includes a polypeptide or enzyme which has at least 60 % amino acid identity as determined by BLAST or FASTA algorithms, preferably at least 75%, most preferably at least 85%, and even more preferably at least 90%, and which has the same or

substantially similar properties or functions as the native or parent protein or enzyme to which it is compared.

As used herein, the term "homologous" in all its grammatical forms and spelling variations refers to the relationship between proteins that possess a "common evolutionary origin," including proteins from superfamilies (*e.g.*, the immunoglobulin superfamily) and homologous proteins from different species (*e.g.*, myosin light chain, etc.) (Reeck *et al.*, Cell 1987, 50:667). Such proteins (and their encoding genes) have sequence homology, as reflected by their sequence similarity, whether in terms of percent similarity or the presence of specific residues or motifs at conserved positions.

Accordingly, the term "sequence similarity" in all its grammatical forms refers to the degree of identity or correspondence between nucleic acid or amino acid sequences of proteins that may or may not share a common evolutionary origin (see Reeck *et al.*, *supra*). However, in common usage and in the instant application, the term "homologous," when modified with an adverb such as "highly," may refer to sequence similarity and may or may not relate to a common evolutionary origin.

In a specific embodiment, two DNA sequences are "substantially homologous" or "substantially similar" when at least about 80%, and most preferably at least about 90 or 95% of the nucleotides match over the defined length of the DNA sequences, as determined by sequence comparison algorithms, such as BLAST, FASTA, DNA Strider, etc. An example of such a sequence is an allelic or species variant of the specific H_4 gene of the invention. Sequences that are substantially homologous can be identified by comparing the sequences using standard software available in sequence data banks, or in a Southern hybridization experiment under, for example, stringent conditions as defined for that particular system.

Similarly, in a particular embodiment, two amino acid sequences are "substantially homologous" or "substantially similar" when greater than 80% of the amino acids are identical, or greater than about 90% are similar (functionally identical). Preferably, the similar or homologous sequences are identified by alignment using, for example, the GCG (Genetics Computer Group, Program Manual for the GCG Package, Version 7, Madison, Wisconsin) pileup program, or any of the programs described above (BLAST, FASTA, etc)

A nucleic acid molecule is "hybridizable" to another nucleic acid molecule, such as a cDNA, genomic DNA, or RNA, when a single stranded form of the nucleic acid molecule

can anneal to the other nucleic acid molecule under the appropriate conditions of temperature and solution ionic strength (see Sambrook *et al.*, *supra*). The conditions of temperature and ionic strength determine the "stringency" of the hybridization. For preliminary screening for homologous nucleic acids, low stringency hybridization conditions, corresponding to a Tm (melting temperature) of 55°C, can be used, *e.g.*, 5x SSC, 0.1% SDS, 0.25% milk, and no formamide; or 30% formamide, 5x SSC, 0.5% SDS. Moderate stringency hybridization conditions correspond to a higher Tm, *e.g.*, 40% formamide, with 5x or 6x SSC. High stringency hybridization conditions correspond to the highest Tm, *e.g.*, 50% formamide, 5x or 6x SSC. SSC is a 0.15M NaCl, 0.015M Na-citrate buffer. Hybridization requires that the two nucleic acids contain complementary sequences, although depending on the stringency of the hybridization, mismatches between bases are possible. The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementation, variables well known in the art. The greater the degree of similarity or homology between two nucleotide sequences, the greater the value of Tm for hybrids of nucleic acids having those sequences. The relative stability (corresponding to higher Tm) of nucleic acid hybridizations decreases in the following order: RNA:RNA, DNA:RNA, DNA:DNA. For hybrids of greater than 100 nucleotides in length, equations for calculating Tm have been derived (see Sambrook *et al.*, *supra*, 9.50-9.51). For hybridization with shorter nucleic acids, *i.e.*, oligonucleotides, the position of mismatches becomes more important, and the length of the oligonucleotide determines its specificity (see Sambrook *et al.*, *supra*, 11.7-11.8). A minimum length for a hybridizable nucleic acid is at least about 10 nucleotides; preferably at least about 15 nucleotides; and more preferably the length is at least about 20 nucleotides.

In a specific embodiment, the term "standard hybridization conditions" refers to a Tm of 55°C, and utilizes conditions as set forth above. In a preferred embodiment, the Tm is 60°C; in a more preferred embodiment, the Tm is 65°C. In a specific embodiment, "high stringency" refers to hybridization and/or washing conditions at 68°C in 0.2xSSC, at 42°C in 50% formamide, 4xSSC, or under conditions that afford levels of hybridization equivalent to those observed under either of these two conditions.

As used herein, the term "oligonucleotide" refers to a nucleic acid, generally of at least 10, preferably at least 15, and more preferably at least 20 nucleotides, preferably no more than 100 nucleotides, that is hybridizable to a genomic DNA molecule, a cDNA molecule, or an

mRNA molecule encoding a gene, mRNA, cDNA, or other nucleic acid of interest. Oligonucleotides can be labeled, *e.g.*, with ^{32}P-nucleotides or nucleotides to which a label, such as biotin, has been covalently conjugated. In one embodiment, a labeled oligonucleotide can be used as a probe to detect the presence of a nucleic acid. In another embodiment, oligonucleotides (one or both of which may be labeled) can be used as PCR primers, either for cloning full length or a fragment of H_4, or to detect the presence of nucleic acids encoding H_4. In a further embodiment, an oligonucleotide of the invention can form a triple helix with a H_4 DNA molecule. Generally, oligonucleotides are prepared synthetically, preferably on a nucleic acid synthesizer. Accordingly, oligonucleotides can be prepared with non-naturally occurring phosphoester analog bonds, such as thioester bonds, etc.

The present invention provides antisense nucleic acids (including ribozymes), which may be used to inhibit expression of H_4 of the invention. Inhibition of H_4 expression may be desired when upregulation of H_4 receptor expression or excessive inhibition of cAMP formation induces disease states such as, transplant organ rejection; asthma; allergies; autoimmune pathologies, such as, but not limited to, multiple sclerosis, type I diabetes, and rheumatoid arthritis; and CNS functions such as cognitive and memory defects. An "antisense nucleic acid" is a single stranded nucleic acid molecule which, on hybridizing under cytoplasmic conditions with complementary bases in an RNA or DNA molecule, inhibits the latter's role. If the RNA is a messenger RNA transcript, the antisense nucleic acid is a countertranscript or mRNA-interfering complementary nucleic acid. As presently used, "antisense" broadly includes RNA-RNA interactions, RNA-DNA interactions, ribozymes and RNase-H mediated arrest. Antisense nucleic acid molecules can be encoded by a recombinant gene for expression in a cell (*e.g.*, U.S. Patent No. 5,814,500; U.S. Patent No. 5,811,234), or alternatively they can be prepared synthetically (*e.g.*, U.S. Patent No. 5,780,607).

Specific non-limiting examples of synthetic oligonucleotides envisioned for this invention include oligonucleotides that contain phosphorothioates, phosphotriesters, methyl phosphonates, short chain alkyl, or cycloalkyl intersugar linkages or short chain heteroatomic or heterocyclic intersugar linkages. Most preferred are those with CH_2-NH-O-CH_2, CH_2-N(CH_3)-O-CH_2, CH_2-O-N(CH_3)-CH_2, CH_2-N(CH_3)-N(CH_3)-CH_2 and O-N(CH_3)-CH_2-CH_2 backbones (where phosphodiester is O-PO_2-O-CH_2). U.S. Patent No. 5,677,437 describes heteroaromatic olignucleoside linkages. Nitrogen linkers or groups containing nitrogen can also be used to

prepare oligonucleotide mimics (U.S. Patents No. 5,792,844 and No. 5,783,682). U.S. Patent No. 5,637,684 describes phosphoramidate and phosphorothioamidate oligomeric compounds. Also envisioned are oligonucleotides having morpholino backbone structures (U.S. Patent No. 5,034,506). In other embodiments, such as the peptide-nucleic acid (PNA) backbone, the phosphodiester backbone of the oligonucleotide may be replaced with a polyamide backbone, the bases being bound directly or indirectly to the aza nitrogen atoms of the polyamide backbone (Nielsen *et al.*, Science 254:1497, 1991). Other synthetic oligonucleotides may contain substituted sugar moieties comprising one of the following at the 2' position: OH, SH, SCH_3, F, OCN, $O(CH_2)_nNH_2$ or $O(CH_2)_nCH_3$ where n is from 1 to about 10; C1 to C10 lower alkyl, substituted lower alkyl, alkaryl or aralkyl; Cl; Br; CN; CF_3; OCF_3; O-; S-, or N-alkyl; O-, S-, or N-alkenyl; $SOCH_3$; SO_2CH_3; ONO_2; NO_2; N_3; NH_2; heterocycloalkyl; heterocycloalkaryl; aminoalkylamino; polyalkylamino; substitued silyl; a fluorescein moiety; an RNA cleaving group; a reporter group; an intercalator; a group for improving the pharmacokinetic properties of an oligonucleotide; or a group for improving the pharmacodynamic properties of an oligonucleotide, and other substituents having similar properties. Oligonucleotides may also have sugar mimetics such as cyclobutyls or other carbocyclics in place of the pentofuranosyl group. Nucleotide units having nucleosides other than adenosine, cytidine, guanosine, thymidine and uridine, such as inosine, may be used in an oligonucleotide molecule.

H_4 Nucleic Acids

A gene encoding H_4, whether genomic DNA or cDNA, can be isolated from any source, particularly from a human cDNA or genomic library. Methods for obtaining H_4 gene are well known in the art, as described above (see, *e.g.*, Sambrook *et al.*, 1989, *supra*). The DNA may be obtained by standard procedures known in the art from cloned DNA (*e.g.*, a DNA "library"), and preferably is obtained from a cDNA library prepared from tissues with high level expression of the protein, by chemical synthesis, by cDNA cloning, or by the cloning of genomic DNA, or fragments thereof, purified from the desired cell (See, for example, Sambrook *et al.*, 1989, *supra*; Glover, D.M. (ed.), 1985, DNA Cloning: A Practical Approach, MRL Press, Ltd., Oxford, U.K. Vol. I, II). Clones derived from genomic DNA may contain regulatory and intron DNA regions in addition to coding regions; clones derived from cDNA will not contain intron sequences. Whatever the source, the gene should be molecularly cloned into a suitable vector for propagation of the gene. Identification of the specific DNA fragment containing the desired

H_4 gene may be accomplished in a number of ways. For example, a portion of a H_4 gene exemplified *infra* can be purified and labeled to prepare a labeled probe, and the generated DNA may be screened by nucleic acid hybridization to the labeled probe (Benton and Davis, Science 1977, 196:180; Grunstein and Hogness, Proc. Natl. Acad. Sci. U.S.A. 1975, 72:3961). Those DNA fragments with substantial homology to the probe, such as an allelic variant from another individual, will hybridize. In a specific embodiment, highest stringency hybridization conditions are used to identify a homologous H_4 gene.

Further selection can be carried out on the basis of the properties of the gene, *e.g.*, if the gene encodes a protein product having the isoelectric, electrophoretic, amino acid composition, partial or complete amino acid sequence, antibody binding activity, or ligand binding profile of H_4 protein as disclosed herein. Thus, the presence of the gene may be detected by assays based on the physical, chemical, immunological, or functional properties of its expressed product.

Other DNA sequences which encode substantially the same amino acid sequence as a H_4 gene may be used in the practice of the present invention. These include but are not limited to allelic variants, species variants, sequence conservative variants, and functional variants.

Amino acid substitutions may also be introduced to substitute an amino acid with a particularly preferable property. For example, a Cys may be introduced a potential site for disulfide bridges with another Cys.

The genes encoding H_4 derivatives and analogs of the invention can be produced by various methods known in the art. The manipulations which result in their production can occur at the gene or protein level. For example, the cloned H_4 gene sequence can be modified by any of numerous strategies known in the art (Sambrook *et al.*, 1989, *supra*). The sequence can be cleaved at appropriate sites with restriction endonuclease(s), followed by further enzymatic modification if desired, isolated, and ligated *in vitro*. In the production of the gene encoding a derivative or analog of H_4, care should be taken to ensure that the modified gene remains within the same translational reading frame as the H_4 gene, uninterrupted by translational stop signals, in the gene region where the desired activity is encoded.

Additionally, the v-encoding nucleic acid sequence can be mutated *in vitro* or *in vivo*, to create and/or destroy translation, initiation, and/or termination sequences, or to create

variations in coding regions and/or form new restriction endonuclease sites or destroy preexisting ones, to facilitate further *in vitro* modification. Such modifications can be made to introduce restriction sites and facilitate cloning the H_4 gene into an expression vector. Any technique for mutagenesis known in the art can be used, including but not limited to, *in vitro* site-directed mutagenesis (Hutchinson, C., *et al.*, J. Biol. Chem. 253:6551, 1978; Zoller and Smith, DNA 3:479-488, 1984; Oliphant *et al.*, Gene 44:177, 1986; Hutchinson *et al.*, Proc. Natl. Acad. Sci. U.S.A. 83:710, 1986), use of TAB™ linkers (Pharmacia), etc. PCR techniques are preferred for site directed mutagenesis (see Higuchi, 1989, "Using PCR to Engineer DNA", in PCR Technology: Principles and Applications for DNA Amplification, H. Erlich, ed., Stockton Press, Chapter 6, pp. 61-70).

The identified and isolated gene can then be inserted into an appropriate cloning vector. A large number of vector-host systems known in the art may be used. Possible vectors include, but are not limited to, plasmids or modified viruses, but the vector system must be compatible with the host cell used. Examples of vectors include, but are not limited to, E. coli, bacteriophages such as lambda derivatives, or plasmids such as Bluescript, pBR322 derivatives or pUC plasmid derivatives, *e.g.*, pGEX vectors, pmal-c, pFLAG, etc. The insertion into a cloning vector can, for example, be accomplished by ligating the DNA fragment into a cloning vector which has complementary cohesive termini. However, if the complementary restriction sites used to fragment the DNA are not present in the cloning vector, the ends of the DNA molecules may be enzymatically modified. Alternatively, any site desired may be produced by ligating nucleotide sequences (linkers) onto the DNA termini; these ligated linkers may comprise specific chemically synthesized oligonucleotides encoding restriction endonuclease recognition sequences. In addition, simple PCR or overlapping PCR may be used to insert a fragment into a cloning vector.

Recombinant molecules can be introduced into host cells via transformation, transfection, infection, electroporation, etc., so that many copies of the gene sequence are generated. Preferably, the cloned gene is contained on a shuttle vector plasmid, which provides for expansion in a cloning cell, *e.g.*, E. coli, and facile purification for subsequent insertion into an appropriate expression cell line, if such is desired. For example, a shuttle vector, which is a vector that can replicate in more than one type of organism, can be prepared for replication in both E. coli and Saccharomyces cerevisiae by linking sequences from an E. coli plasmid with

sequences form the yeast 2μ plasmid.

H_4 Regulatory Nucleic Acids

Elements of the H_4 promoter can be identified by scanning the human genomic region upstream of the H_4 start site, *e.g.*, by creating deletion mutants and checking for expression, or with the TRANSFAC algorithm. Sequences up to about 6 kilobases (kb) or more upstream from the H_4 start site can contain tissue-specific regulatory elements.

The term "H_4 promoter" encompasses artificial promoters. Such promoters can be prepared by deleting nonessential intervening sequences from the upstream region of the H_4 promoter, or by joining upstream regulatory elements from the H_4 promoter with a heterologous minimal promoter, such as the CMV immediate early promoter.

An H_4 promoter can be operably associated with a heterogenous coding sequence, *e.g.*, for reporter gene (luciferase and green fluorescent proteins are examples of reporter genes) in a construct. This construct will result in expression of the heterologous coding sequence under control the H_4 promoter, *e.g.*, a reporter gene can be expressed, under conditions that under normal conditions cause H_4 expression. This construct can be used in screening assays, described below, for H_4 agonists and antagonists.

Expression of H_4 Polypeptides

The nucleotide sequence coding for H_4, or antigenic fragment, derivative or analog thereof, or a functionally active derivative, including a chimeric protein, thereof, can be inserted into an appropriate expression vector, *i.e.*, a vector which contains the necessary elements for the transcription and translation of the inserted protein-coding sequence. Thus, a nucleic acid encoding H_4 of the invention can be operationally associated with a promoter in an expression vector of the invention. Both cDNA and genomic sequences can be cloned and expressed under control of such regulatory sequences. Such vectors can be used to express functional or functionally inactivated H_4 polypeptides.

The necessary transcriptional and translational signals can be provided on a recombinant expression vector, or they may be supplied by the native gene encoding H_4 and/or its flanking regions.

Potential host-vector systems include but are not limited to mammalian cell systems transfected with expression plasmids or infected with virus (*e.g.*, vaccinia virus, adenovirus, adeno-associated virus, herpes virus, etc.); insect cell systems infected with virus

(*e.g.*, baculovirus); microorganisms such as yeast containing yeast vectors; or bacteria transformed with bacteriophage, DNA, plasmid DNA, or cosmid DNA. The expression elements of vectors vary in their strengths and specificities. Depending on the host-vector system utilized, any one of a number of suitable transcription and translation elements may be used.

Expression of H_4 protein may be controlled by any promoter/enhancer element known in the art, but these regulatory elements must be functional in the host selected for expression. Promoters which may be used to control H_4 gene expression include, but are not limited to, cytomegalovirus (CMV) promoter (U.S. Patent Nos. 5,385,839 and 5,168,062), the SV40 early promoter region (Benoist and Chambon, 1981, Nature 290:304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto, *et al.*, Cell 22:787-797, 1980), the herpes thymidine kinase promoter (Wagner *et al.*, Proc. Natl. Acad. Sci. U.S.A. 78:1441-1445, 1981), the regulatory sequences of the metallothionein gene (Brinster *et al.*, Nature 296:39-42, 1982); prokaryotic expression vectors such as the β-lactamase promoter (Villa-Komaroff, *et al.*, Proc. Natl. Acad. Sci. U.S.A. 75:3727-3731, 1978), or the tac promoter (DeBoer, *et al.*, Proc. Natl. Acad. Sci. U.S.A. 80:21-25, 1983); see also "Useful proteins from recombinant bacteria" in Scientific American, 242:74-94, 1980; promoter elements from yeast or other fungi such as the Gal 4 promoter, the ADC (alcohol dehydrogenase) promoter, PGK (phosphoglycerol kinase) promoter, alkaline phosphatase promoter; and transcriptional control regions that exhibit tissue specificity, particularly endothelial cell-specific promoters.

Solubilized forms of the protein can be obtained by solubilizing inclusion bodies or reconstituting membrane components, *e.g.*, by treatment with detergent, and if desired sonication or other mechanical processes, as described above. The solubilized protein can be isolated using various techniques, such as polyacrylamide gel electrophoresis (PAGE), isoelectric focusing, 2-dimensional gel electrophoresis, chromatography (*e.g.*, ion exchange, affinity, immunoaffinity, and sizing column chromatography), centrifugation, differential solubility, immunoprecipitation, or by any other standard technique for the purification of proteins.

Vectors

A wide variety of host/expression vector combinations may be employed in expressing the DNA sequences of this invention. Useful expression vectors, for example, may consist of segments of chromosomal, non-chromosomal and synthetic DNA sequences. Suitable vectors include derivatives of SV40 and known bacterial plasmids, *e.g.*, E. coli plasmids col El,

pCR1, pBR322, pMal-C2, pET, pGEX (Smith *et al.*, Gene 67:31-40, 1988), pMB9 and their derivatives, plasmids such as RP4; phage DNAS, *e.g.*, the numerous derivatives of phage l, *e.g.*, NM989, and other phage DNA, *e.g.*, M13 and filamentous single stranded phage DNA; yeast plasmids such as the 2μ plasmid or derivatives thereof; vectors useful in eukaryotic cells, such as vectors useful in insect or mammalian cells; vectors derived from combinations of plasmids and phage DNAs, such as plasmids that have been modified to employ phage DNA or other expression control sequences; and the like.

Viral vectors, such as lentiviruses, retroviruses, herpes viruses, adenoviruses, adeno-associated viruses, vaccinia virus, baculovirus, alphavirus, and other recombinant viruses with desirable cellular tropism are also useful. Thus, a gene encoding a functional or mutant H_4 protein or polypeptide domain fragment thereof can be introduced *in vivo*, *ex vivo*, or *in vitro* using a viral vector or through direct introduction of DNA. Expression in targeted tissues can be effected by targeting the transgenic vector to specific cells, such as with a viral vector or a receptor ligand, or by using a tissue-specific promoter, or both. Targeted gene delivery is described in International Patent Publication WO 95/28494, published October 1995.

Viral vectors commonly used for *in vivo* or *ex vivo* targeting and therapy procedures are DNA-based vectors and retroviral vectors. Methods for constructing and using viral vectors are known in the art (see, *e.g.*, Miller and Rosman, BioTechniques 1992, 7:980-990). Preferably, the viral vectors are replication defective, that is, they are unable to replicate autonomously in the target cell. In general, the genome of the replication defective viral vectors which are used within the scope of the present invention lack at least one region which is necessary for the replication of the virus in the infected cell. These regions can either be eliminated (in whole or in part) or be rendered non-functional by any technique known to a person skilled in the art. These techniques include the total removal, substitution (by other sequences, in particular by the inserted nucleic acid), partial deletion or addition of one or more bases to an essential (for replication) region. Such techniques may be performed *in vitro* (on the isolated DNA) or *in situ*, using the techniques of genetic manipulation or by treatment with mutagenic agents. Preferably, the replication defective virus retains the sequences of its genome which are necessary for encapsidating the viral particles.

DNA viral vectors include an attenuated or defective DNA virus, such as but not limited to herpes simplex virus (HSV), papillomavirus, Epstein Barr virus (EBV), adenovirus,

adeno-associated virus (AAV), and the like. Defective viruses, which entirely or almost entirely lack viral genes, are preferred. Defective virus is not infective after introduction into a cell. Use of defective viral vectors allows for administration to cells in a specific, localized area, without concern that the vector can infect other cells. Thus, a specific tissue can be specifically targeted. Examples of particular vectors include, but are not limited to, a defective herpes virus 1 (HSV1) vector (Kaplitt *et al.*, Molec. Cell. Neurosci. 2:320-330, 1991), defective herpes virus vector lacking a glyco-protein L gene (Patent Publication RD 371005 A), or other defective herpes virus vectors (International Patent Publication No. WO 94/21807, published September 29, 1994; International Patent Publication No. WO 92/05263, published April 2, 1994); an attenuated adenovirus vector, such as the vector described by Stratford-Perricaudet *et al.* (J. Clin. Invest. 90:626-630, 1992; see also La Salle *et al.*, Science 259:988-990, 1993); and a defective adeno-associated virus vector (Samulski *et al.*, J. Virol. 61:3096-3101, 1987; Samulski *et al.*, J. Virol. 63:3822-3828, 1989; Lebkowski *et al.*, Mol. Cell. Biol. 8:3988-3996, 1988).

Various companies produce viral vectors commercially, including but by no means limited to Avigen, Inc. (Alameda, CA; AAV vectors), Cell Genesys (Foster City, CA; retroviral, adenoviral, AAV vectors, and lentiviral vectors), Clontech (retroviral and baculoviral vectors), Genovo, Inc. (Sharon Hill, PA; adenoviral and AAV vectors), Genvec (adenoviral vectors), IntroGene (Leiden, Netherlands; adenoviral vectors), Molecular Medicine (retroviral, adenoviral, AAV, and herpes viral vectors), Norgen (adenoviral vectors), Oxford BioMedica (Oxford, United Kingdom; lentiviral vectors), and Transgene (Strasbourg, France; adenoviral, vaccinia, retroviral, and lentiviral vectors).

Preferably, for *in vivo* administration, an appropriate immunosuppressive treatment is employed in conjunction with the viral vector, *e.g.*, adenovirus vector, to avoid immuno-deactivation of the viral vector and transfected cells. For example, immunosuppressive cytokines, such as interleukin-12 (IL-12), interferon-γ (IFN-γ), or anti-CD4 antibody, can be provided to block humoral or cellular immune responses to the viral vectors (see, *e.g.*, Wilson, Nature Medicine, 1995). In that regard, it is advantageous to employ a viral vector that is engineered to express a minimal number of antigens.

In another embodiment, the vector can be introduced *in vivo* by lipofection, as naked DNA, or with other transfection facilitating agents (peptides, polymers, etc.). Synthetic cationic lipids can be used to prepare liposomes for *in vivo* transfection of a gene encoding a

marker (Felgner, et. al., Proc. Natl. Acad. Sci. U.S.A. 84:7413-7417, 1987; Felgner and Ringold, Science 337:387-388, 1989; see Mackey, *et al.*, Proc. Natl. Acad. Sci. U.S.A. 85:8027-8031, 1988; Ulmer, *et al.*, Science 259:1745-1748, 1993). Useful lipid compounds and compositions for transfer of nucleic acids are described in International Patent Publications WO 95/18863 and WO 96/17823, and in U.S. Patent No. 5,459,127. Lipids may be chemically coupled to other molecules for the purpose of targeting (see Mackey, *et al.*, *supra*). Targeted peptides, *e.g.*, hormones or neurotransmitters, and proteins such as antibodies, or non-peptide molecules could be coupled to liposomes chemically.

Other molecules are also useful for facilitating transfection of a nucleic acid *in vivo*, such as a cationic oligopeptide (*e.g.*, International Patent Publication WO 95/21931), peptides derived from DNA binding proteins (*e.g.*, International Patent Publication WO 96/25508), or a cationic polymer (*e.g.*, International Patent Publication WO95/21931).

Alternatively, non-viral DNA vectors for gene therapy can be introduced into the desired host cells by methods known in the art, *e.g.*, electroporation, microinjection, cell fusion, DEAE dextran, calcium phosphate precipitation, use of a gene gun (ballistic transfection; see, *e.g.*, U.S. Pat. No. 5,204,253, U.S. Pat. No. 5,853,663, U.S. Pat. No. 5,885,795, and U.S. Pat. No. 5,702,384 and see Sanford, TIB-TECH, 6:299-302, 1988; Fynan *et al.*, Proc. Natl. Acad. Sci. U.S.A., 90:11478-11482, 1993; and Yang *et al.*, Proc. Natl. Acad. Sci. U.S.A., 87:1568-9572, 1990), or use of a DNA vector transporter (see, *e.g.*, Wu, *et al.*, J. Biol. Chem. 267:963-967, 1992; Wu and Wu, J. Biol. Chem. 263:14621-14624, 1988; Hartmut, *et al.*, Canadian Patent Application No. 2,012,311, filed March 15, 1990; Williams, *et al.*, Proc. Natl. Acad. Sci. USA 88:2726-2730, 1991). Receptor-mediated DNA delivery approaches can also be used (Curiel, *et al.*, Hum. Gene Ther. 3:147-154, 1992; Wu and Wu, J. Biol. Chem. 262:4429-4432, 1987). U.S. Patent Nos. 5,580,859 and 5,589,466 disclose delivery of exogenous DNA sequences, free of transfection facilitating agents, in a mammal. Recently, a relatively low voltage, high efficiency *in vivo* DNA transfer technique, termed electrotransfer, has been described (Mir, *et al.*, C.P. Acad. Sci., 321:893, 1998; WO 99/01157; WO 99/01158; WO 99/01175).

H_4 Ligands and Binding Partners

The present invention further permits identification of physiological ligands and binding partners of H_4. One method for evaluating and identifying H_4 binding partners is the

yeast two-hybrid screen. Preferably, the yeast two-hybrid screen is performed using an cell library with yeast that are transformed with recombinant H_4. Alternatively, H_4 can be used as a capture or affinity purification reagent. In another alternative, labeled H_4 can be used as a probe for binding, *e.g.*, by immunoprecipitation or Western analysis. Expected H_4 binding partners are G-proteins.

Generally, binding interactions between H_4 and any of its binding partners will be strongest under conditions approximating those found in the cytoplasm, *i.e.*, physiological conditions of ionic strength, pH and temperature. Perturbation of these conditions will tend to disrupt the stability of a binding interaction.

Antibodies to H_4

Antibodies to H_4 are useful, *inter alia*, for diagnostics and intracellular regulation of H_4 activity, as set forth below. According to the invention, a H_4 polypeptide produced recombinantly or by chemical synthesis, and fragments or other derivatives or analogs thereof, including fusion proteins, may be used as immunogens to generate antibodies that recognize the H_4 polypeptide. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments, and an Fab expression library. Such an antibody is preferfably specific for human H_4 and it may recognize either a mutant form of H_4 or wild-type H_4, or both.

One can use the hydropathic index of amino acids, as discussed by Kyte and Doolittle (J Mol Biol. 1982, 157:105-132). See, for example, U.S. Patent 4,554,101, which states that the greatest local average hydrophilicity of a "protein," as governed by the hydrophilicity of its adjacent amino acids, correlates with its immunogenicity. Accordingly, it is noted that substitutions can be made based on the hydrophilicity assigned to each amino acid. In using either the hydrophilicity index or hydropathic index, which assigns values to each amino acid, it is preferred to introduce substitutions of amino acids where these values are ± 2, with ± 1 being particularly preferred, and those within ± 0.5 being the most preferred substitutions.

Various procedures known in the art may be used for the production of polyclonal antibodies to H_4 polypeptide or derivative or analog thereof. For the production of antibody, various host animals can be immunized by injection with the H_4 polypeptide, or a derivative (*e.g.*, fragment or fusion protein) thereof, including but not limited to rabbits, mice, rats, sheep, goats, etc. In one embodiment, the H_4 polypeptide or fragment thereof can be conjugated to an

immunogenic carrier, *e.g.*, bovine serum albumin (BSA) or keyhole limpet hemocyanin (KLH). Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and Corynebacterium parvum.

For preparation of monoclonal antibodies directed toward the H_4 polypeptide, or fragment, analog, or derivative thereof, any technique that provides for the production of antibody molecules by continuous cell lines in culture may be used. These include but are not limited to the hybridoma technique originally developed by Kohler and Milstein (Nature 1975, 256:495-497), as well as the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, Immunology Today 1983, 4:72; Cote *et al.*, Proc. Natl. Acad. Sci. 1983, 80:2026-2030), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole *et al.*, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., 1985, pp. 77-96). In an additional embodiment of the invention, monoclonal antibodies can be produced in germ-free animals (International Patent Publication No. WO 89/12690). In fact, according to the invention, techniques developed for the production of "chimeric antibodies" (Morrison *et al.*, J. Bacteriol. 1984, 159:870; Neuberger *et al.*, Nature 1984, 312:604-608; Takeda *et al.*, Nature 1985, 314:452-454) by splicing the genes from a mouse antibody molecule specific for an H_4 polypeptide together with genes from a human antibody molecule of appropriate biological activity can be used; such antibodies are within the scope of this invention. Such human or humanized chimeric antibodies are preferred for use in therapy of human diseases or disorders (described *infra*), since the human or humanized antibodies are much less likely than xenogenic antibodies to induce an immune response, in particular an allergic response, themselves.

Antibody fragments which contain the idiotype of the antibody molecule can be generated by known techniques. For example, such fragments include but are not limited to: the F(ab')2 fragment which can be produced by pepsin digestion of the antibody molecule; the Fab' fragments which can be generated by reducing the disulfide bridges of the F(ab')2 fragment, and the Fab fragments which can be generated by treating the antibody molecule with papain and a reducing agent.

According to the invention, techniques described for the production of single

chain antibodies (U.S. Patent Nos. 5,476,786, 5,132,405, and U.S. Patent 4,946,778) can be adapted to produce H_4 polypeptide-specific single chain antibodies. An additional embodiment of the invention utilizes the techniques described for the construction of Fab expression libraries (Huse *et al.*, Science 1989, 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity for an H_4 polypeptide, or its derivatives, or analogs.

In the production and use of antibodies, screening for or testing with the desired antibody can be accomplished by techniques known in the art, *e.g.*, radioimmunoassay, ELISA (enzyme-linked immunosorbant assay), "sandwich" immunoassays, immunoradiometric assays, gel diffusion precipitin reactions, immunodiffusion assays, *in situ* immunoassays (using colloidal gold, enzyme or radioisotope labels, for example), Western blots, precipitation reactions, agglutination assays (*e.g.*, gel agglutination assays, hemagglutination assays), complement fixation assays, immunofluorescence assays, protein A assays, and immunoelectrophoresis assays, etc. In one embodiment, antibody binding is detected by detecting a label on the primary antibody. In another embodiment, the primary antibody is detected by detecting binding of a secondary antibody or reagent to the primary antibody. In a further embodiment, the secondary antibody is labeled. Many means are known in the art for detecting binding in an immunoassay and are within the scope of the present invention. For example, to select antibodies which recognize a specific epitope of an H_4 polypeptide, one may assay generated hybridomas for a product which binds to an H_4 polypeptide fragment containing such epitope. For selection of an antibody specific to an H_4 polypeptide from a particular species of animal, one can select on the basis of positive binding with H_4 polypeptide expressed by or isolated from cells of that species of animal.

The foregoing antibodies can be used in methods known in the art relating to the localization and activity of the H_4 polypeptide, *e.g.*, for Western blotting, imaging H_4 polypeptide *in situ*, measuring levels thereof in appropriate physiological samples, etc. using any of the detection techniques mentioned above or known in the art. Such antibodies can also be used in assays for ligand binding, *e.g.*, as described in U.S. Patent No. 5,679,582. Antibody binding generally occurs most readily under physiological conditions, *e.g.*, pH of between about 7 and 8, and physiological ionic strength. The presence of a carrier protein in the buffer solutions stabilizes the assays. While there is some tolerance of perturbation of optimal conditions, *e.g.*, increasing or decreasing ionic strength, temperature, or pH, or adding detergents or chaotropic

salts, such perturbations will decrease binding stability.

In a specific embodiment, antibodies that act as ligands and agonize or antagonize the activity of H_4 polypeptide can be generated. In addition, intracellular single chain Fv antibodies can be used to regulate cAMP formation (Marasco et al., Proc. Natl. Acad. Sci. U.S.A. 1993, 90:7884-7893; Chen., Mol. Med. Today 1997, 3:160-167; Spitz et al., Anticancer Res. 1996, 16:3415-22; Indolfi et al., Nat. Med. 1996, 2:634-635; Kijma et al., Pharmacol. Ther. 1995, 68:247-267). Such antibodies can be tested using the assays described *infra* for identifying ligands.

Screening and Chemistry

According to the present invention, nucleotide sequences encoding H_4 and the H_4 receptor structure, which can be modeled from the amino acid sequence based on homology to other GPCR proteins, are useful targets to identify drugs that are effective in treating disorders associated with histamine-regulated processes. Drug targets include without limitation (i) isolated nucleic acids derived from the gene encoding H_4 (*e.g.*, antisense or ribozyme molecules) and (ii)small molecule compounds that recognize and bind the receptor.

In particular, identification and isolation of H_4 provides for development of screening assays, particularly for high throughput screening of molecules that up- or down-regulate the activity of H_4. Accordingly, the present invention contemplates methods for identifying specific histamine receptor ligands that interact with H_4 receptors, using various screening assays known in the art.

Any screening technique known in the art can be used to screen for H_4 agonists or antagonists. The present invention contemplates screens for small molecule ligands or ligand analogs and mimics, as well as screens for natural ligands that bind to and agonize or antagonize H_4 activity *in vivo*. For example, natural products libraries can be screened using assays of the invention for molecules that agonize or antagonize H_4 expression or activity.

Another approach uses recombinant bacteriophage to produce large libraries. Using the "phage method" (Scott and Smith, Science 1990, 249:386-390; Cwirla, *et al.*, Proc. Natl. Acad. Sci., USA 1990, 87:6378-6382; Devlin *et al.*, Science 1990, 49:404-406), very large libraries can be constructed (106-108 chemical entities). A second approach uses primarily chemical methods, of which the Geysen method (Geysen *et al.*, Molecular Immunology 1986,

23:709-715; Geysen *et al.* J. Immunologic Method 1987 102:259-274; and the method of Fodor *et al.* (Science 1991, 251:767-773) are examples. Furka *et al.* (14th International Congress of Biochemistry, Volume #5 1988, Abstract FR:013; Furka, Int. J. Peptide Protein Res. 1991, 37:487-493), Houghton (U.S. Patent No. 4,631,211) and Rutter (U.S. Patent No. 5,010,175) describe methods to produce a mixture of peptides that can be tested as agonists or antagonists.

In another aspect, synthetic libraries (Needels *et al.*, Proc. Natl. Acad. Sci. USA 1993, 90:10700-4; Ohlmeyer *et al.*, Proc. Natl. Acad. Sci. USA 1993, 90:10922-10926; Lam *et al.*, PCT Publication No. WO 92/00252; Kocis *et al.*, PCT Publication No. WO 9428028) and the like can be used to screen for ligands that regulate H_4 activity. Test compounds are screened from large libraries of synthetic or natural compounds. Numerous means are currently used for random and directed synthesis of saccharide, peptide, and nucleic acid based compounds. Synthetic compound libraries are commercially available from Maybridge Chemical Co. (Trevillet, Cornwall, UK), Comgenex (Princeton, NJ), Brandon Associates (Merrimack, NH), and Microsource (New Milford, CT). A rare chemical library is available from Aldrich (Milwaukee, WI). Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available from *e.g.* Pan Laboratories (Bothell, WA) or MycoSearch (NC), or are readily producible. Additionally, natural and synthetically produced libraries and compounds are readily modified through conventional chemical, physical, and biochemical means (Blondelle *et al.*, Tib Tech 1996, 14:60).

Knowledge of the primary sequence of H_4, and the similarity of that sequence with proteins of known function, can provide an initial clue as to the structure of agonists or antagonists of the receptor. Identification and screening of agonists antagonists is further facilitated by determining structural features of the receptor, *e.g.*, using X-ray crystallography, neutron diffraction, nuclear magnetic resonance spectrometry, homology studies, structure-activity relationships, and other techniques for structure determination. These techniques provide for the rational design or identification of agonists and antagonists.

One technique that may be used to assess the affinity of a test compound for the H_4 receptor is a competition binding assay. In this assay, test wells containing an aliquot of a lipid bilayer membranes that contain the histamine H_4 receptor are incubated with an known concentration of a radiolabeled ligand for the receptor. The lipid bilayer may be prepared by any known protocol that separates the membrane containing receptor component from the

cytoplasmic components. Each well also is incubated with a different concentration of a unlabeled test compound. Cell membranes are then separated from the incubation mixture by any method known in the art including, but not limited to, centrifugation and vacuum filtration on a cell harvester. The radioactivity of each well is then determined using any device that can detect radioactivity, such as a scintillation counter. As increasing concentrations of the test compound compete for the receptor binding site, the radioactivity detected decreases. The data then can be converted using the Cheng-Prusoff equation (Biochem Pharmacol. 1973, 22:3099-3108) to determine the affinity (K_i) of the compound for the receptor.

In vivo screening methods

Intact cells or whole animals expressing a gene encoding H_4 can be used in screening methods to identify candidate drugs.

In one series of embodiments, a permanent cell line is established. Alternatively, cells (including without limitation mammalian, insect, yeast, or bacterial cells) are transiently programmed to express an H_4 gene by introduction of appropriate DNA or mRNA. Identification of candidate compounds can be achieved using any suitable assay, including without limitation (i) assays that measure binding of test compounds to H_4 (ii) assays that measure the ability of a test compound to modify (*i.e.*, inhibit or enhance) a measurable activity or function of H_4 and (iii) assays that measure the ability of a compound to modify (*i.e.*, inhibit or enhance) the transcriptional activity of sequences derived from the promoter (*i.e.*, regulatory) regions of the H_4 gene.

H_4 knockout mammals can be prepared for evaluating the molecular pathology of this defect in greater detail than is possible with human subjects. Such animals also provide excellent models for screening drug candidates. A "knockout mammal" is an mammal (*e.g.*, mouse, rabbit) that contains within its genome a specific gene that has been inactivated by the method of gene targeting (see, *e.g.*, U.S. Patent Nos. 5,777,195 and 5,616,491). A knockout mammal includes both a heterozygote knockout (*i.e.*, one defective allele and one wild-type allele) and a homozygous mutant (*i.e.*, two defective alleles; a heterologous construct for expression of an H_4, such as a human H_4, could be inserted to permit the knockout mammal to live if lack of H_4 expression was lethal). Preparation of a knockout mammal requires first introducing a nucleic acid construct that will be used to suppress expression of a particular gene into an undifferentiated cell type termed an embryonic stem cell. This cell is then injected into

a mammalian embryo. A mammalian embryo with an integrated cell is then implanted into a foster mother for the duration of gestation. Zhou, *et al.* (Genes and Development 1995, 9:2623-34) describes PPCA knock-out mice.

The term "knockout" refers to partial or complete suppression of the expression of at least a portion of a protein encoded by an endogenous DNA sequence in a cell. The term "knockout construct" refers to a nucleic acid sequence that is designed to decrease or suppress expression of a protein encoded by endogenous DNA sequences in a cell. The nucleic acid sequence used as the knockout construct is typically comprised of (1) DNA from some portion of the gene (exon sequence, intron sequence, and/or promoter sequence) to be suppressed and (2) a marker sequence used to detect the presence of the knockout construct in the cell. The knockout construct is inserted into a cell, and integrates with the genomic DNA of the cell in such a position so as to prevent or interrupt transcription of the native DNA sequence. Such insertion usually occurs by homologous recombination (*i.e.*, regions of the knockout construct that are homologous to endogenous DNA sequences hybridize to each other when the knockout construct is inserted into the cell and recombine so that the knockout construct is incorporated into the corresponding position of the endogenous DNA). The knockout construct nucleic acid sequence may comprise (1) a full or partial sequence of one or more exons and/or introns of the gene to be suppressed, (2) a full or partial promoter sequence of the gene to be suppressed, or (3) combinations thereof. Typically, the knockout construct is inserted into an embryonic stem cell (ES cell) and is integrated into the ES cell genomic DNA, usually by the process of homologous recombination. This ES cell is then injected into, and integrates with, the developing embryo.

The phrases "disruption of the gene" and "gene disruption" refer to insertion of a nucleic acid sequence into one region of the native DNA sequence (usually one or more exons) and/or the promoter region of a gene so as to decrease or prevent expression of that gene in the cell as compared to the wild-type or naturally occurring sequence of the gene. By way of example, a nucleic acid construct can be prepared containing a DNA sequence encoding an antibiotic resistance gene which is inserted into the DNA sequence that is complementary to the DNA sequence (promoter and/or coding region) to be disrupted. When this nucleic acid construct is then transfected into a cell, the construct will integrate into the genomic DNA. Thus, many progeny of the cell will no longer express the gene at least in some cells, or will express it at a

decreased level, as the DNA is now disrupted by the antibiotic resistance gene.

Generally, the DNA will be at least about 1 kb in length and preferably 3-4 kb in length, thereby providing sufficient complementary sequence for recombination when the knockout construct is introduced into the genomic DNA of the ES cell (discussed below).

Included within the scope of this invention is a mammal in which two or more genes have been knocked out. Such mammals can be generated by repeating the procedures set forth herein for generating each knockout construct, or by breeding to mammals, each with a single gene knocked out, to each other, and screening for those with the double knockout genotype.

Regulated knockout animals can be prepared using various systems, such as the tet-repressor system (see U.S. Patent No. 5,654,168) or the Cre-Lox system (see U.S. Patent Nos. 4,959,317 and 5,801,030).

In another series of embodiments, transgenic animals are created in which (i) a human H_4 is stably inserted into the genome of the transgenic animal; and/or (ii) the endogenous H_4 genes are inactivated and replaced with human H_4 genes. See, *e.g.*, Coffman, Semin. Nephrol. 1997, 17:404; Esther *et al.*, Lab. Invest. 1996, 74:953; Murakami et al., Blood Press. Suppl. 1996, 2:36.

H_4 Activation Assay

Any cell assay system that allows for assessment of functional activity of H_4 agonists and antagonists is defined by the present invention. In a specific embodiment, exemplified *infra*, the assay can be used to identify compounds that selectively interact with H_4, which can be evaluated by assessing the effects of H_4 transformed cells contacted with a test compound, which modulates cAMP accumulation. The assay system can thus be used to identify compounds that selectively produce a functional effect through histamine H_4 receptors. Compounds that increase cAMP formation and accumulation may be useful as novel therapeutics in the prevention of transplanted organ rejection, asthma, allergies, and autoimmune pathologies, such as, but not limited to, multiple sclerosis, type I diabetes, and rheumatoid arthritis. Additionally, CNS functions such as cognitive and memory defects may be treated. Preferably, each experiment is performed in triplicate at multiple different dilutions of test compound.

An agonist and/or antagonist screen involves detecting cAMP accumulation by the host cell when contacted with H_4 ligand. If cAMP accumulation is increased, the test

compound is a candidate antagonist of H_4 receptors. If cAMP accumulation is decreased, the test compound is a candidate agonist of H_4 receptors. If there is no change in cAMP formation, the test compound is not an effective H_4 ligand.

Any convenient method permits detection of the formed product, cAMP. For example, the invention provides immunoassays for detecting cAMP. Typically, immunoassays use either a labeled antibody or a labeled antigenic component (*e.g.*, that competes with the antigen in the sample for binding to the antibody). Suitable labels include without limitation enzyme-based, fluorescent, chemiluminescent, radioactive, or dye molecules. Assays that amplify the signals from the probe are also known, such as, for example, those that utilize biotin and avidin, and enzyme-labelled immunoassays, such as ELISA assays. Alternatively, labeled antigenic component may be quantified by scintillation techniques. In another method, the second messenger, preferably cAMP, will be separated on a high performance liquid chromatograph and quantified by a UV detector.

The assay system described here also may be used in a high-throughput primary screen for agonists and antagonists, or it may be used as a secondary functional screen for candidate compounds identified by a different primary screen, *e.g.*, a binding assay screen that identifies compounds that interact with the receptor.

High-Throughput Screen

Agents according to the invention may be identified by screening in high-throughput assays, including without limitation cell-based or cell-free assays. It will be appreciated by those skilled in the art that different types of assays can be used to detect different types of agents. Several methods of automated assays have been developed in recent years so as to permit screening of tens of thousands of compounds in a short period of time. Such high-throughput screening methods are particularly preferred. The use of high-throughput screening assays to test for agents is greatly facilitated by the availability of large amounts of purified polypeptides, as provided by the invention.

Compounds

"Histamine" refers to a neurotransmitter that is produced and released from neurons. Histamine is formed from the amino acid histidine by histidine decarboxylase. Structurally, histamine is an imidazolethylamine. In other words, histamine is comprised of

an imidazole moiety and an amino group connected by an alkyl chain. The pharmacologically active form of histamine is proposed to be the monocationic tautomer, where one of the nitrogen present in the imidazole ring is positively charged. However, different forms of histamine may interact with histamine receptors to produce a functional effect. Histamine is produced intracellularly and stored until released in response to a physiological stimulus.

"Histamine analogs" refers to compounds that comprise the imidazole, alkyl chain, and a nitrogen moiety (*e.g.*, amine (unsubstituted or substituted), piperidine, pyridine) of histamine, but may be modified at other positions. These modifications may be performed to alter affinity and/or selectivity of the compound for the histamine receptors. "Histamine compounds" refers to compounds that may bind to the histamine receptors.

Histamine analogs and compounds can be classified as agonists or antagonists. As discussed previously, agonists are ligands that bind to the receptor and produce a functional effect similar to that produced by the endogenous ligand (*i.e.*, histamine) for the receptor, whereas antagonists are ligands that bind to the receptor and block a functional effect produced by an agonist for the receptor or the endogenous ligand of the receptor. Histamine analogs and compounds are further described in Goodman and Gilman's The Pharmacological Basis of Therapeutics, Ninth Edition, McGraw-Hill, 1996.

Agonists that may be contemplated by this invention include, but are not limited to, R-(α)-methylhistamine, imetit, and immepip. Antagonists, burimamide, impromidine, dimaprit, and thioperamide clobenpropit and iodophenpropit impentamine, GT2016 and iodoproxyfan. Other compounds include derivatives, metabolites, and precursors.

Methods of Diagnosis

According to the present invention, genetic variants of H_4 can be detected to diagnose an H_4 associated disease, such as treatment for transplanted organ rejection, asthma, allergies, and autoimmune pathologies, such as, but not limited to, multiple sclerosis, type I diabetes, and rheumatoid arthritis. Additionally, CNS functions such as cognitive and memory defects may be treated. The various methods for detecting such variants are described herein. Where such variants impact H_4 function, either as a result of a mutated

amino acid sequence or because the mutation results in expression of a truncated protein, or no expression at all, they are expected to result in disregulation of the allergic response, the immune response, cognition and memory.

Nucleic Acid Assays

The DNA may be obtained from any cell source. DNA is extracted from the cell source or body fluid using any of the numerous methods that are standard in the art. It will be understood that the particular method used to extract DNA will depend on the nature of the source. Generally, the minimum amount of DNA to be extracted for use in the present invention is about 25 pg (corresponding to about 5 cell equivalents of a genome size of 4 x 10^9 base pairs).

In another alternate embodiment, RNA is isolated from biopsy tissue using standard methods well known to those of ordinary skill in the art such as guanidium thiocyanate-phenol-chloroform extraction (Chomocyznski *et al.*, Anal. Biochem., 162:156, 1987). The isolated RNA is then subjected to coupled reverse transcription and amplification by polymerase chain reaction (RT-PCR), using specific oligonucleotide primers that are specific for a selected site. Conditions for primer annealing are chosen to ensure specific reverse transcription and amplification; thus, the appearance of an amplification product is diagnostic of the presence of a particular genetic variation. In another embodiment, RNA is reverse-transcribed and amplified, after which the amplified sequences are identified by, *e.g.*, direct sequencing. In still another embodiment, cDNA obtained from the RNA can be cloned and sequenced to identify a mutation.

Protein Assays

In an alternate embodiment, biopsy tissue is obtained from a subject. Antibodies that are capable of specifically binding to H_4 are then contacted with samples of the tissue to determine the presence or absence of a H_4 polypeptide specified by the antibody. The antibodies may be polyclonal or monoclonal, preferably monoclonal. Measurement of specific antibody binding to cells may be accomplished by any known method, *e.g.*, quantitative flow cytometry, enzyme-linked or fluorescence-linked immunoassay, Western analysis, etc.

Therapeutic Uses

According to the present invention, stimulation of H_4 receptor activity may be used as a treatment option in patients with histamine-related disease states. Stimulation of H_4 receptor activity may be by methods, such as, but not limited to, (i) providing polypeptides that stimulate receptor activity and (ii) providing compounds that stimulate receptor activity.

Gene Therapy

In a specific embodiment, vectors comprising a sequence encoding a protein, including, but not limited to, full-length H_4, are provided to treat or prevent a disease or disorder associated with the function of H_4 in peripheral blood leukocytes. In this embodiment of the invention, the therapeutic vector encodes a sequence that produces the protein of the invention.

Any of the methods for gene therapy available in the art can be used according to the present invention. Exemplary methods are described below.

For general reviews of the methods of gene therapy, see, Goldspiel *et al.*, Clinical Pharmacy, 1993, 12:488-505; Wu and Wu, Biotherapy, 1991, 3:87-95; Tolstoshev, Ann. Rev. Pharmacol. Toxicol., 1993, 32:573-596; Mulligan, Science, 1993, 260:926-932; and Morgan and Anderson, Ann. Rev. Biochem., 1993, 62:191-217; May, TIBTECH, 1993, 11:155-215. Methods commonly known in the art of recombinant DNA technology that can be used are described in Ausubel *et al.*, (eds.), 1993, Current Protocols in Molecular Biology, John Wiley & Sons, NY; Kriegler, 1990, Gene Transfer and Expression, A Laboratory Manual, Stockton Press, NY; and in Chapters 12 and 13, Dracopoli et al., (eds.), 1994, Current Protocols in Human Genetics, John Wiley & Sons, NY. Vectors suitable for gene therapy are described above.

In one aspect, the therapeutic vector comprises a nucleic acid that expresses a protein of the invention in a suitable host. In particular, such a vector has a promoter operationally linked to the coding sequence for the protein. The promoter can be inducible or constitutive and, optionally, tissue-specific. In another embodiment, a nucleic acid molecule is used in which the protein coding sequences and any other desired sequences are flanked by regions that promote homologous recombination at a desired site in the genome, thus providing for intrachromosomal expression of the protein (Koller and Smithies, Proc. Natl. Acad. Sci. U.S.A, 1989, 86:8932-8935; Zijlstra *et al.*, Nature, 1989, 342:435-438).

Delivery of the vector into a patient may be either direct, in which case the patient is directly exposed to the vector or a delivery complex, or indirect, in which case, cells are first

transformed with the vector *in vitro* then transplanted into the patient. These two approaches are known, respectively, as *in vivo* and *ex vivo* gene therapy.

In a specific embodiment, the vector is directly provided *in vivo*, where it enters the cells of the organism and mediates expression of the protein. This can be accomplished by any of numerous methods known in the art, by constructing it as part of an appropriate expression vector and administering it so that it becomes intracellular, *e.g.*, by infection using a defective or attenuated retroviral or other viral vector (see, U.S. Patent No. 4,980,286), or by direct injection of naked DNA, or by use of microparticle bombardment (*e.g.*, a gene gun; Biolistic, Dupont); or coating with lipids or cell-surface receptors or transfecting agents, encapsulation in biopolymers (*e.g.*, poly-S-1-64-N-acetylglucosamine polysaccharide; see, U.S. Patent No. 5,635,493), encapsulation in liposomes, microparticles, or microcapsules; by administering it in linkage to a peptide or other ligand known to enter the nucleus; or by administering it in linkage to a ligand subject to receptor-mediated endocytosis (see, *e.g.*, Wu and Wu, J. Biol. Chem., 1987, 62:4429-4432), etc. In another embodiment, a nucleic acid ligand complex can be formed in which the ligand comprises a fusogenic viral peptide to disrupt endosomes, allowing the nucleic acid to avoid lysosomal degradation. In yet another embodiment, the nucleic acid can be targeted *in vivo* for cell specific uptake and expression, by targeting a specific receptor (see, *e.g.*, PCT Publication Nos. WO 92/06180, WO 92/22635, WO 92/20316 and WO 93/14188). Alternatively, the nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression by homologous recombination (Koller and Smithies, Proc. Natl. Acad. Sci. USA, 1989, 86:8932-8935; Zijlstra, et al., Nature, 1989, 342:435-438). These methods are in addition to those discussed above in conjunction with "Viral and Non-viral Vectors".

The form and amount of therapeutic nucleic acid envisioned for use depends on the type of disease and the severity of the desired effect, patient state, etc., and can be determined by one skilled in the art.

Inhibition or stimulation of protein synthesis

Gene transcription and protein translation may be inhibited or stimulated by administration of exogenous compounds. Exogenous compounds may interact with extracellular and/or intracellular messenger systems, such as, but not limited to, adenosine triphosphate, nitric oxide, and guanosine triphosphate; to regulate protein synthesis. In this embodiment, exogenous compounds that stimulate or inhibit H_4 protein synthesis may be

used in the prevention and/or treatment for transplanted organ rejection, asthma, allergies, and autoimmune pathologies, such as, but not limited to, multiple sclerosis, type I diabetes, and rheumatoid arthritis. Additionally, CNS functions such as cognitive and memory defects may be treated.

The present invention provides antisense nucleic acids (including ribozymes), which may be used to inhibit expression of H_4 of the invention. The antisense nucleic acid, upon hybridizing under cytoplasmic conditions with complementary bases in an RNA or DNA molecule, inhibits the role of the RNA or DNA. Additionally, hybridization of the antisense nucleic acid to the DNA or RNA may inhibit transcription of the DNA into RNA and/or translation of the RNA into the protein. If the RNA is a messenger RNA transcript, the antisense nucleic acid is a countertranscript or mRNA-interfering complementary nucleic acid. Antisense nucleic acid molecules can be encoded by a recombinant gene for expression in a cell (*e.g.*, U.S. Patent No. 5,814,500; U.S. Patent No. 5,811,234) or can be prepared synthetically (*e.g.*, U.S. Patent No. 5,780,607).

Alternatively, antibody molecules can also be administered, for example, by expressing nucleotide sequences encoding single-chain antibodies within the target cell population by utilizing, for example, techniques such as those described in Marasco *et al.* (Proc. Natl. Acad Sci. USA, 1993, 90:7889-7893).

Therapeutically suggested compounds may be provided to the patient in formulations that are known in the art and may include any pharmaceutically acceptable additives, such as excipents, lubricants, diluents, flavorants, colorants, and disintegrants. The formulations may be produced in useful dosage units such as tablet, caplet, capsule, liquid, or injection.

The form and amount of therapeutic compound envisioned for use depends on the type of disease and the severity of the desired effect, patient state, etc., and can be determined by one skilled in the art.

EXAMPLES

The present invention will be better understood by reference to the following Examples, which are provided as exemplary of the invention, and not by way of limitation.

EXAMPLE I: CHARACTERIZATION OF THE H_4 RECEPTOR

A GPCR sequence profile was generated from a sequence alignment of members of this subfamily based on hidden Markov Models (Eddy Bioinformatics, 1998 14:755-763; Durbin *et al.* A tutorial introduction to hidden Markov models and other probabilistic modeling approaches in computational sequence analysis, Cambridge University Press, 1998) to look for novel members in the human genome database. In detail, GPCR proteins of the biogenic amine subfamily were retrieved from Swiss-Prot database and a sequence alignment was generated by a multiple sequence alignment tool, named CLUSTALW (Thompson *et al.* Nucleic Acids Res, 1994, 22:4673-80). Using the HMMER program (Eddy, HMMER User's Guide and Program, Version 2.1, 1998), a consensus sequence (GNLLVILVIL RTKKLRTPTN IFILNLAVAD LLFLLTLPPW ALYYLVGGSE DWPFGSALCK LVTALDVVNM YASILLLTAI SIDRYLAIVH PLRYRRRRTS PRRAKVVILL VWVLALLLSL PPLLFSWVKT VEEGNGTLNV NVTVCLIDFP EESTASVSTW LRSYVLLSTL VGFLLPLLVI LVCYTRILRT LRKAAKTLLV VVVVFVLCWL PYFIVLLLDT LCLSIIMSST CELERVLPTA LLVTLWLAYV NSCLNPIIY; SEQ ID NO: 3) was developed from the biogenic amine subfamily members. The consensus sequence contained the unique 7-transmembrane sequence structure of biogenic amine GPCRs.

A weekly update of nucleotide sequence from GenBank database is maintained in-house. An auto-search script using TBLASTN program was written and the biogenic amine GPCR consensus sequence was used to search this database weekly. Every TBLASTN search result was carefully examined and potential open reading frame (ORF) fragments were extracted from nucleotide sequence. Each fragment was further verified to determine its novelty.

Four peptide fragments translated from a recently released human genomic sequence (Accession number: AC007922) of chromosome 18 clone RP11-178F10, from the Whitehead Institute/MIT Center for Genome Research, were shown to have modest homology to different regions of this GPCR consensus sequence. These four fragments located to different regions of sequence AC009722 in both plus and minus strands. Interestingly, the highest scoring hit from searching the protein database with these peptides by BLASP was in all cases, the human histamine receptor 3 (H_3). It is very likely that these four fragments are exons of a GPCR gene and their appearance in both strands of this genomic sequence may result from incorrect genomic contig assembly.

To clone the full length ORF, multiple primer sets were designed against the predicted sequence. Using multiple primer pairs, 3 overlapping pieces spanning the entire ORF were obtained from a human heart marathon cDNA library (Clontech, Palo Alto, CA). Using the marathon adapter primer API, and a reverse primer corresponding to ORF bases 644-615, a fragment of the H_4 cDNA was obtained which corresponds to bases 276-644 of the H_4 coding region (exon sequence) and 39 bases of intronic sequence at the 5' end. Primers comprising nucleotides 461-482 and 1173-1146 were used to PCR a 712bp band. The 5' sequence was obtained using primers comprising nucleotides 1-32 and 339-309. The outer primers (1-32, 1173-1146) were used to piece the 3 fragments together. The sequence was ligated into the mammalian expression vector pCDNA3.1 + zeo (Invitrogen, Carlsbad, CA).

EXAMPLE II: TISSUE EXPRESSION OF THE H_4 RECEPTOR

Quantitative RT-PCR was performed on an ABI 7700 "Taqman" sequence detection system to determine the tissue distribution of the H_4 receptor. Primers spanning the exon 1-2 boundary (Forward primer: 5'-taacttggccattgacttctt-3' (SEQ ID NO:4); Reverse primer: 5'-attcgaacagcatgtgagggat-3' (SEQ ID NO:5) and a Probe: 5'-(6-carboxyfluorescein)-tacaaaggaatggagatcacacccaca-(6-carboxy-N,N,N'N'-tetramethylrhodamine)-3' (SEQ ID NO:6)) were used to determine H_4 expression levels in a mRNA prepared from a series of human tissues, purchased from Clontech (Palo Alto). A 2 step reaction procedure was performed as per manufacturers directions. Briefly, 2mg RNA was reverse transcribed using random hexamers (2.5mM in a final volume of 20ml. 14ul of this was used in the PCR reaction. The cyling conditions were as follows, 95 °C for 10 minutes, followed by 40 cycles of (a) 95 °C for 15 seconds and (b) 60 °C for 1 minute.

EXAMPLE III: PHARMACOLOGICAL PROFILE OF THE H_4 RECEPTOR IN YEAST CELLS

The H_4 receptor protein coding sequences were amplified using Forward oligo: 5'-aaggatccaaaatgccagatactaatagc-3' (SEQ ID NO:7) and Reverse oligo: 5'-aagtcgacttaagaagatactgaccgac-3' (SEQ ID NO:8) that add BamHI and yeast consensus translational initiation sites to the 5' end and a SalI site to the 3' end. The fragment was cloned into corresponding sites in the multicopy yeast expression vector, p426GPD, thus placing

receptor expression under control of the strong constitutive GPD1 promoter. The yeast expression plasmid, pMP327, was introduced into MPY578i5 cells (MATa ura3 his3 trp1 leu2 lys2 ade2 far1::LYS2 fus1::FUS1-HIS3 sst2::SST2-G418R ste2::LEU2, gpa1::GPA1i5) (Hadcock, J.R. and Pausch, M.H. submitted) using LiOAc and selected for ura prototrophy. In order to facilitate coupling of the receptor to the G protein, MPY578i5 cells express a chimeric G alpha protein coupled to the mating signal transduction pathway. The chimeric construct is expressed from the GPA1 locus and is composed of Gpa1 sequences in which the 5'-C-terminal amino acids have been replaced with those of Gai3. A multicopy FUS1-LacZ reporter gene plasmid, pMP283 (Hadcock, J.R. and Pausch, M.H. submitted), was subsequently introduced into H_4R-containing MPY578i5 cells and selected on media lacking trp and ura. The resulting yeast strain, MPY733, was used for further analysis.

Samples (250 ng) of compounds present in the LOPAC panel (Sigma RBI, Natick MA). were dispensed to 96 well microtiter dishes. MPY733 cells (5 x 10^5/ml, 200 ml/well) in assay medium (SCD-ura-trp, pH 6.8, 25 mM PIPES, 0.1mg/ml Chlorophenylred bglactopyranoside(CPRG), 2 mM 3-AT) were added and cultured overnight at 30 °C. The presence of active compounds was detected the next day by measurement of absorbance at 570 nm using a Wallac Victor II. The LOPAC panel was screened in duplicate and in parallel with another yeast strain containing a different orphan GPCR. Only compounds that produced significantly elevated absorbance in both receptor containing plates and not in the other GPCR containing plates were deemed active.

EXAMPLE IV: PHARMACOLOGICAL PROFILE OF THE H_4 RECEPTOR IN MAMMALIAN CELLS

The ORF was modified by PCR for mammalian expression of H_4. A 5' HindIII restriciton enzymes site and a Kozak consensus sequence were added using the primer 5'-aagcttccaccatgccagatactaatagcacaatcaatttatc-3' (SEQ ID NO:9), and a 3' Xba1 site added 5'-tctagattaagaagatactgaccgactgtgttg-3' (SEQ ID NO:10). The sequence was confirmed and ligated into the HindIII and Xba I sites of the mammalian expression vector pCDNA3.1+ zeo (Invitrogen, Carlsbad CA). HEK 293 cells (approximately 10^7 cells) were transfected with the pCDNA3.1+zeo/H_4 using standard lipofectamine plus reagent (Life Technologies). Cells were maintained in DMEM containing 10% fetal calf serum and penicillin (100units/ml)/streptomycin

(100mg/ml). 48 hours post transfection cells were selected using 500mg/ml zeocin. Zeocin resistant clones were assayed using the cAMP assay and by RT-PCR. RNA was extracted from approximately 10^6 cells using One step PCR kit (Life technologies). The primers used in the extraction were 5'-ggaaggatgaaggtagtgaatg-3' (SEQ ID NO:11) and 5'-cagaatctgattgggaggaagg-3' (SEQ ID NO:12).

HEK cells stably expressing the H_4 receptor were assessed functionally in cAMP assays using the cAMP scintillation proximity assay (SPA) (Amersham Pharmacia Biotech, Pistcataway, NJ). Briefly, 40 000 cells were plated into wells of a 96 well plate. 24 hours later the media was removed and replaced with 100µl Krebs bicarbonate buffer and the cells were incubated at 37°C for 15 minutes. Following this, the cells were incubated in Krebs buffer containing 0.5mM isobutylmethylxanthine, a phosphodiesterase inhibitor, to prevent cAMP breakdown. The effect of H_4 on the forskolin induced formation was determined by incubating the cells in the presence of forskolin (10µM) and agonist for 12 minutes. cAMP levels were determined using the cAMP SPA kit according to the manufacturers directions.

EXAMPLE V: H_4 EXPRESSION IN INDUCED CD4 AND CD8 CLONES

This examination was undertaken using the TaqMan EZ RT PCR kit and the following oligonucleotides (provided by Philip Jones at Wyeth Neuroscience) (i) H_4 EX 1F (forward oligo): 5'-taacttggccatctttgac-3' (SEQ ID NO:13), (ii) H_4 EX 1R (reverse primer): 5'-attcgaacagcgtgtgag-3' (SEQ ID NO:14), and H_4 EX 1 Probe: 5'-(6-carboxyfluorescein)-tacaaaggaatggagatca-3' (SEQ ID NO:15). RNA for the standard curve was polyA+ human leukocyte RNA from Clontech.

50 nanograms of total RNA was assayed in duplicate from 6 human T cell clones derived from a single human donor. Three of these clones were CD4+, two were CD8+, and one was an NK T cell clone. These T cell clones were stimulated with anti-CD3 and RNA was isolated at 0, 2, 4, 8, 24 and 48 hours after stimulation. Expression was normalized with GAPDH.

In order to determine whether the pattern of H_4 expression was specific to clonal populations of human lymphocytes, the same TaqMan quantitative PCR assay on 10 nanograms of RNA from bulk populations of CD4+ and CD8+ T cells isolated from three different individuals. These lymphocytes were stimulated with anti-CD3 and harvested at the same time

points listed above. H_4 expression was normalized with GAPDH.

RESULTS

Using a GPCR sequence profile generated from a Hidden Markov Model (HMM) of the biogenic amine subfamily, several sequences with homology to the human H3 receptor were identified from the human genomic sequence of chromosome 18 (clone RP11- 178F10 Accession number: AC007922 from the Whitehead Institute/MIT Center for Genome Research) (Eddy Bioinformatics, 1998, 14:755-763). A contig of the predicted exons forms a sequence encoding a putative GPCR whose nucleotide sequence and translated peptide sequence are shown in Figure 1 and SEQ ID NOS:1 and 2. The sequence was deposited with Genbank (Accession Number AF307973).

The predicted protein sequence contains residues that are conserved across the biogenic amine receptors. As in other biogenic amine receptors, the present protein contains conserved aspartate residues within transmembrane (TM) domains 2 and 3 (at positions 61 and 94). A DY (D94 and Y95) motif in TM3 is found only in histamine and muscarinic receptors. A DRY sequence motif that exists at the cytoplasmic interface of TM3 of all GPCRs also can be found in the H_4 sequence (amino acids 111- 113). Also present are two conserved cysteine residues in the first and second extracellular loops (position 88 and 164), which are predicted to form a disulfide bond; two conserved tryptophan residues in TM4 and TM6 (positions 140 and 316); and conserved proline residues in TM5, TM6, and TM7 (positions 186, 318, and 355, respectively). Based on these results, H_4 cDNA appears to encode a biogenic amine-like receptor.

The predicted open reading frame is 1173 base pairs long and encodes a protein of 390 amino acids. Sequence comparison using BLASTP, under standard conditions, analysis reveals that the novel protein is most similar to the human H_3 receptor (44% identical and 51% similar). Based on sequence homology it is proposed that the receptor belongs to the histamine receptor family, therefore we have termed it the H_4 receptor.

Receptor distribution studies indicate that the present receptor is highly expressed in peripheral blood leukocytes. Trace amounts of the H_4 receptor are expressed in heart, lung and placenta. It is proposed that these trace amounts of expression represent blood cell mRNA present in the samples.

Functional studies further confirmed the categorization of the protein as a novel histamine receptor. Histamine and the H_3 selective agonist R-α-methylhistamine stimulated cAMP accumulation, suggesting these compounds are agonists at the H_4 receptor (See Figure 2A). R-α-methylhistamine exhibited both lower potency and efficacy than histamine at the H_4 receptor despite being more potent at the related cloned H_3 receptor and in several tissue based assays for the H_3 receptor (Vollinga, *et al.*, J. Med. Chem. 1995, 38:266-271; Harper, *et al.*, Br. J. Pharmacol. 1999, 128:751-759). The selective H_3 receptor antagonist clobenpropit was a partial agonist at the H_4 receptor (See Figure 2A).

The H_3 antagonist thioperamide almost fully inhibited the stimulatory response produced by histamine (See Figure 2B). Comparatively, clobenpropit partially blocked histamine-induced stimulation of cAMP formation, further suggesting that clobenpropit is a partial agonist at the H_4 receptor. The interaction of the H_4 receptor to a GPAI/ Gαi3 chimeric G-protein alpha subunit also predicts its coupling specificity in mammalian cells.

Expression of the H_4 receptor in HEK 293tsa cells confirms the coupling of the H_4 receptor to the inhibition of cAMP formation (See Figure 2C). In these cells, forskolin stimulated 8-fold greater cAMP formation compared to basal levels. Addition of histamine (1 μM) inhibited forskolin-induced stimulation of cAMP accumulation by about 40%. These studies suggest that the H_4 receptor couples an inhibitory G-protein to inhibit adenylyl cyclase activity and cAMP accumulation.

Structure-activity relationship studies were conducted with several histamine antagonists to further define the pharmacological profile of this receptor. Studies were conducted to correlate the effect of the alkyl chain to antagonist activity at the histamine H_4 receptor (See Figure 3). The rank order of efficacy obtained for the human H_4 was (number represents the length of the alkyl chain) 5>6= 1 0>8=4. This rank order of efficacy contrasts with the human H_3 receptor where the rank order is 5>4>6>8>10.

Prior studies have indicated the expression of at least two subtypes of the H_3 receptor. These two potential H_3 subtypes, localized in the rat brain and guinea pig jejenum, have been shown to have rank order of efficiencies of 4=5>3>6>8 and 5>6=4>8>3, respectively. These assays were different (rat was a radioligand binding assay and guinea pig was an organ bath experiment using isolated guinea pig ileum). However, the relative potencies of the series can be compared, so its possible to say that as the rank order differs then the receptors are likely

to be different (Leurs *et al.* J. Pharm. Exp. Ther., 1996, 276:1009-1015.)

The present studies also indicate that thioperamide is more efficacious and potent than impentamine at the H_4 receptor. Comparatively, impentamine has greater affinity for cloned histamine H_3 receptors (4-fold) than thioperamide (Ki=50.8nM and 193nM respectively).

H_4 expression was consistently detected in all three CD4+ clones at most time points (Figure 4). There was no consistent temporal pattern to anti-CD3 induction of expression between these three clones. One CD8+ clone showed a very low but detectable level of expression at most time points, while the other as well as the NKT clone had no detectable H_4 expression despite good GAPDH amplification.

H_4 expression, normalized with GAPDH, was detectable in all samples at all time points (Figure 5). Anti-CD3 stimulation produced significant induction of H_4 expression in all 6 samples, most peaking at the 8 hour time point. The highest level of induction was found in the 8 hour CD4+ and CD8+ samples from the same donor (designated 'U'), peaking at greater than 100 copies of H_4 RNA per copy of GAPDH.

Combined, these studies indicate that the pharmacological profile of the H_4 receptor is not similar to any known histamine H_3 receptor.

* * *

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and the accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

It is further to be understood that values are approximate, and are provided for description.

Patents, patent applications, publications, procedures, and the like are cited throughout this application, the disclosures of which are incorporated herein by reference in their entireties.

WHAT IS CLAIMED IS:

1. An isolated H_4 receptor protein having an amino acid sequence at least 51% identical or comprising at least 10 contiguous amino acids from the sequence depicted in SEQ ID NO:2 which H_4 receptor protein binds ligands comprising an imidazole and an amine, which imidazole and amine are attached by an alkyl chain, wherein the rank order of efficacy of modulation second messenger formation of the ligands at the H_4 receptor protein is 5>6=10>8=4, where the number represents the number of carbons in the alkyl chain.

2. The H_4 receptor protein of claim 1, wherein upon binding histamine or a histamine agonist the receptor protein inhibits second messenger formation.

3. The H_4 receptor protein of claim 2, wherein the second messenger is cAMP.

4. The H_4 receptor protein of claim 1 which is a human H_4 receptor protein.

5. The H_4 receptor protein of claim 4 which has an amino acid sequence as depicted in SEQ ID NO: 2.

6. The H_4 receptor protein of claim 4 which is encoded by a nucleic acid having a sequence as depicted in SEQ ID NO: 1.

7. An isolated H_4 receptor protein having an amino acid sequence with at least 95% sequence identity to human H_4 receptor protein having an amino acid sequence as depicted in SEQ ID NO: 2.

8. An isolated nucleic acid encoding the H_4 receptor protein of claim 1 or 7.

9. The nucleic acid of claim 8 which is a cDNA.

10. The nucleic acid of claim 8, wherein the H_4 receptor protein is a human H_4 receptor protein.

11. An isolated nucleic acid encoding an H_4 receptor protein, which nucleic acid hybridizes under stringent conditions to a nucleic acid having a sequence of at least 20 nucleotides identical to a corresponding nucleotide sequence of the same number of bases in SEQ ID NO:1 or its complement.

12. The nucleic acid of claim 8, which encodes an H_4 receptor protein having an amino acid sequence as depicted in SEQ ID NO:2.

13. The nucleic acid of claim 12, which comprises a nucleotide sequence as depicted in SEQ ID NO:1.

14. A vector comprising the nucleic acid of claim 8 operably associated with an expression control sequence.

15. A host cell transfected with the vector of claim 14.

16. A non-human animal transformed with the vector of claim 14, wherein the animal expresses a H_4 receptor protein at a detectable level, whereby the cells expressing the H_4 receptor protein suppress cAMP formation when contacted with an H_4 receptor agonist.

17. A method for producing a H_4 receptor protein, which method comprises culturing host cells of claim 15 under conditions that provide for expression of the H_4 receptor protein by the vector.

18. An isolated nucleic acid of at least ten bases having a nucleotide sequence identical to a corresponding nucleotide sequence of the same number of bases in SEQ ID NO: 1 or its complement.

19. The nucleic acid of claim 18 which is detectably labeled.

20. An antibody that specifically binds to the H_4 receptor protein of claim 1 or 7.

21. A method for detecting an H_4 receptor protein, which method comprises detecting binding of the antibody of claim 20 to a protein in a sample suspected of containing a H_4 receptor protein, wherein the antibody is contacted with the sample under conditions that permit specific binding with any H_4 receptor protein present in the sample.

22. A method for detecting expression of H_4 receptor, which method comprises detecting mRNA encoding H_4 receptor in a sample from a cell suspected of expressing H_4 receptor.

23. The method according to claim 22 wherein mRNA encoding H_4 receptor is detected by hybridization to a H_4 receptor-specific nucleic acid.

24. The method according to claim 23 wherein the H_4 receptor-specific nucleic acid is at least 10 nucleotides in length and has a sequence identical to a sequence of the same number of bases in SEQ ID NO: 1, or the complementary sequence thereof.

25. An assay system for identifying H_4 receptor ligands, comprising a sufficient number of cells of claim 15 to detect an alteration in second messenger accumulation.

26. The assay system of claim 25, wherein the second messenger is cAMP.

27. The assay system of claim 25, wherein the receptor is a human receptor.

28. A method for identifying a test compound that antagonizes histamine H_4 receptors, which method comprises detecting an increase in the level of a second messenger in an assay system of claim 25 contacted with the test compound, wherein an increase in the level of the second messenger indicates that the test compound antagonizes the H_4 receptor.

29. A method for identifying a test compound that agonizes histamine H_4 receptors, which method comprises detecting a decrease in the level of a second messenger in an assay system of claim 25 contacted with the test compound, wherein the decrease in the level of the second messenger indicates that the test compound agonizes the H_4 receptor.

30. A method for identifying a compound that binds an H_4 receptor, which method comprises detecting binding of a test compound to the H_4 receptor protein of claim 1.

31. The method according to claim 29, wherein binding of the test compound is detecting by inhibiting binding of a labeled H_4 ligand.

32. The method according to claim 29, wherein the H_4 receptor protein is in a lipid bilayer membrane.

33. An isolated nucleic acid that specifically hybridizes under highly stringent conditions to the complement of the sequence depicted in SEQ ID NO:1, wherein said nucleic acid encodes a histamine H_4 receptor protein.

FIG. 1A

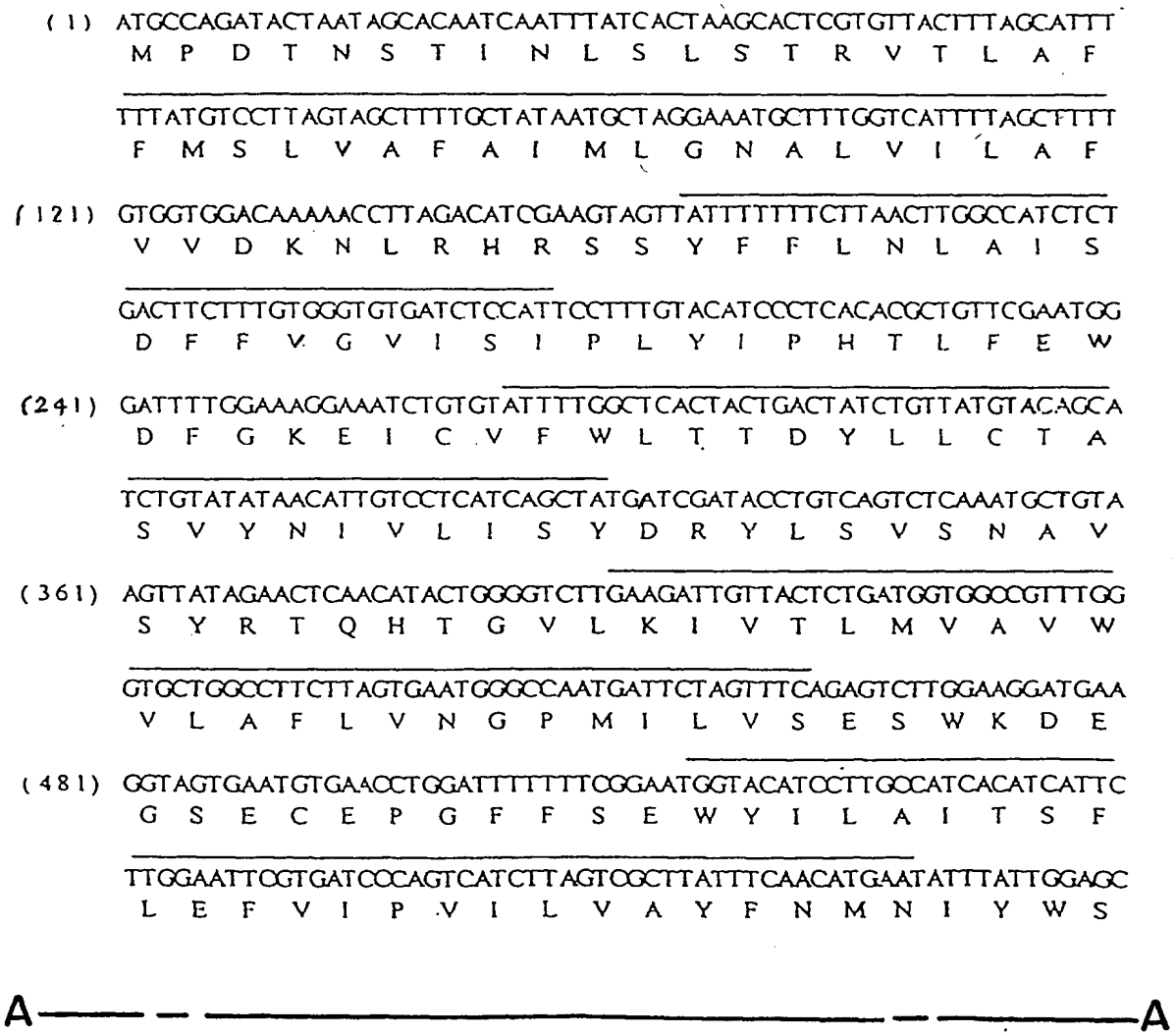

(1) ATGCCAGATACTAATAGCACAATCAATTTATCACTAAGCACTCGTGTTACTTTAGCATTT
M P D T N S T I N L S L S T R V T L A F
TTTATGTCCTTAGTAGCTTTTGCTATAATGCTAGGAAATGCTTTGGTCATTTTAGCTTTT
F M S L V A F A I M L G N A L V I L A F
(121) GTGGTGGACAAAAACCTTAGACATCGAAGTAGTTATTTTTTTCTTAACTTGGCCATCTCT
V V D K N L R H R S S Y F F L N L A I S
GACTTCTTTGTGGGTGTGATCTCCATTCCTTTGTACATCCCTCACACGCTGTTCGAATGG
D F F V G V I S I P L Y I P H T L F E W
(241) GATTTTGGAAAGGAAATCTGTGTATTTTGGCTCACTACTGACTATCTGTTATGTACAGCA
D F G K E I C V F W L T T D Y L L C T A
TCTGTATATAACATTGTCCTCATCAGCTATGATCGATACCTGTCAGTCTCAAATGCTGTA
S V Y N I V L I S Y D R Y L S V S N A V
(361) AGTTATAGAACTCAACATACTGGGGTCTTGAAGATTGTTACTCTGATGGTGGCCGTTTGG
S Y R T Q H T G V L K I V T L M V A V W
GTGCTGGCCTTCTTAGTGAATGGGCCAATGATTCTAGTTTCAGAGTCTTGGAAGGATGAA
V L A F L V N G P M I L V S E S W K D E
(481) GGTAGTGAATGTGAACCTGGATTTTTTTCGGAATGGTACATCCTTGCCATCACATCATTC
G S E C E P G F F S E W Y I L A I T S F
TTGGAATTCGTGATCCCAGTCATCTTAGTCGCTTATTTCAACATGAATATTTATTGGAGC
L E F V I P V I L V A Y F N M N I Y W S
A
A

FIG. 1B

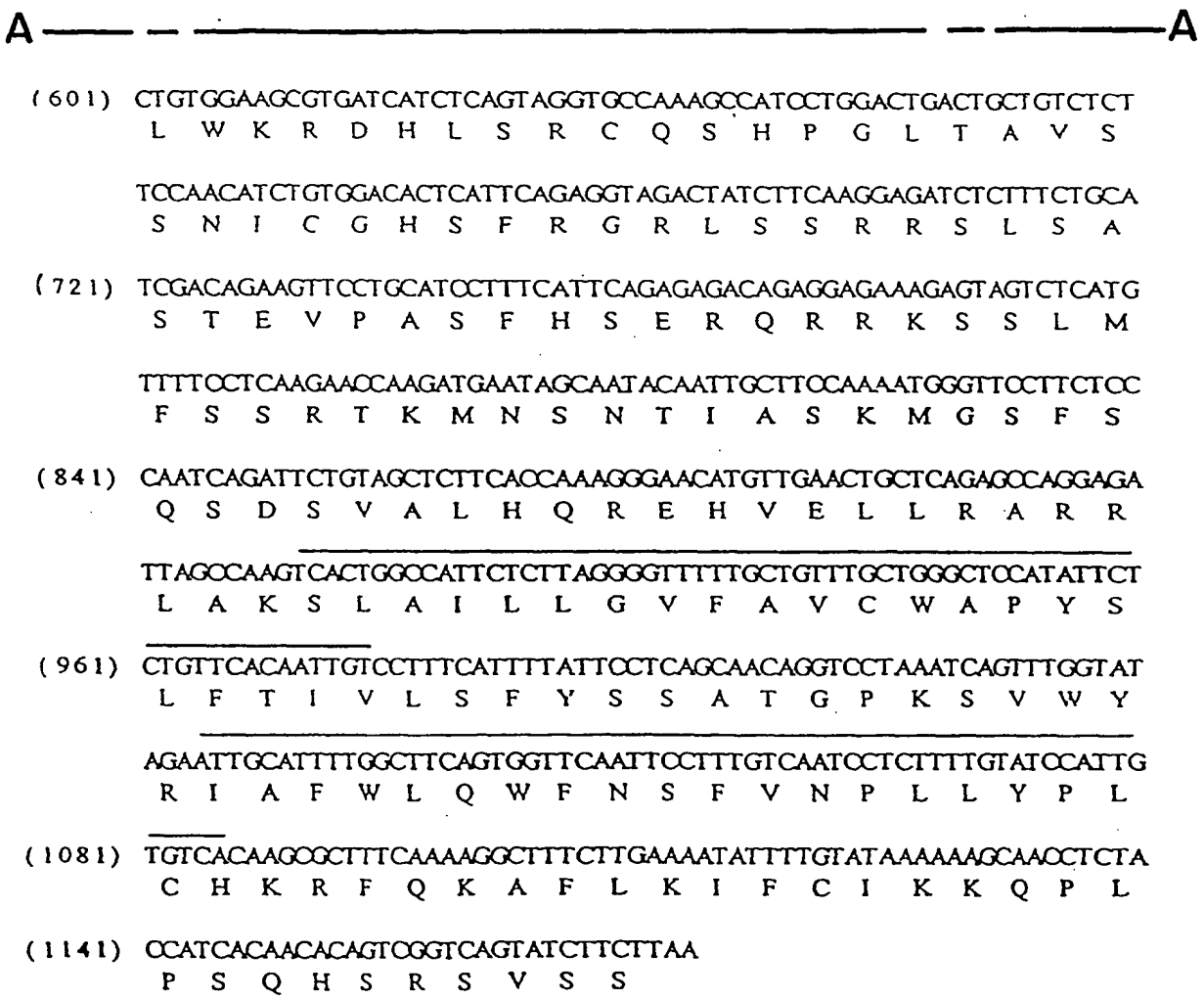

A
A
(601) CTGTGGAAGCGTGATCATCTCAGTAGGTGCCAAAGCCATCCTGGACTGACTGCTGTCTCT
L W K R D H L S R C Q S H P G L T A V S
TCCAACATCTGTGGACACTCATTCAGAGGTAGACTATCTTCAAGGAGATCTCTTTCTGCA
S N I C G H S F R G R L S S R R S L S A
(721) TCGACAGAAGTTCCTGCATCCTTTCATTCAGAGAGACAGAGGAGAAAGAGTAGTCTCATG
S T E V P A S F H S E R Q R R K S S L M
TTTTCCTCAAGAACCAAGATGAATAGCAATACAATTGCTTCCAAAATGGGTTCCTTCTCC
F S S R T K M N S N T I A S K M G S F S
(841) CAATCAGATTCTGTAGCTCTTCACCAAAGGGAACATGTTGAACTGCTCAGAGCCAGGAGA
Q S D S V A L H Q R E H V E L L R A R R
TTAGCCAAGTCACTGGCCATTCTCTTAGGGGTTTTTGCTGTTTGCTGGGCTCCATATTCT
L A K S L A I L L G V F A V C W A P Y S
(961) CTGTTCACAATTGTCCTTTCATTTTATTCCTCAGCAACAGGTCCTAAATCAGTTTGGTAT
L F T I V L S F Y S S A T G P K S V W Y
AGAATTGCATTTTGGCTTCAGTGGTTCAATTCCTTTGTCAATCCTCTTTTGTATCCATTG
R I A F W L Q W F N S F V N P L L Y P L
(1081) TGTCACAAGCGCTTTCAAAAGGCTTTCTTGAAAATATTTTGTATAAAAAAGCAACCTCTA
C H K R F Q K A F L K I F C I K K Q P L
(1141) CCATCACAACACAGTCGGTCAGTATCTTCTTAA
P S Q H S R S V S S

FIG. 2A

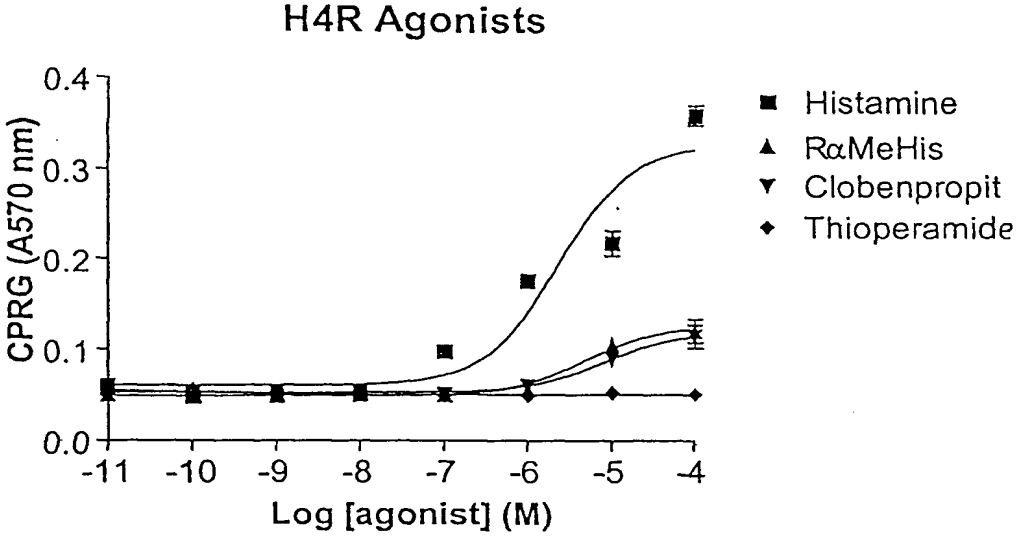

H4R Agonists
CPRG (A570 nm)
0.4
0.3
0.2
0.1
0.0
-11 -10 -9 -8 -7 -6 -5 -4
Log [agonist] (M)
Histamine
RαMeHis
Clobenpropit
Thioperamide

FIG. 2B

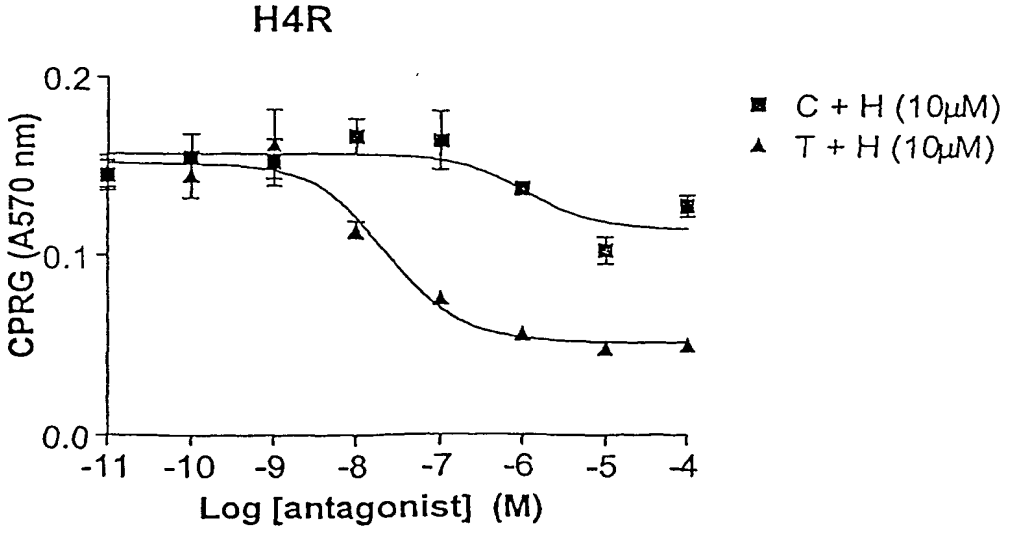

H4R
CPRG (A570 nm)
0.2
0.1
0.0
-11 -10 -9 -8 -7 -6 -5 -4
Log [antagonist] (M)
C + H (10μM)
T + H (10μM)

FIG. 2C

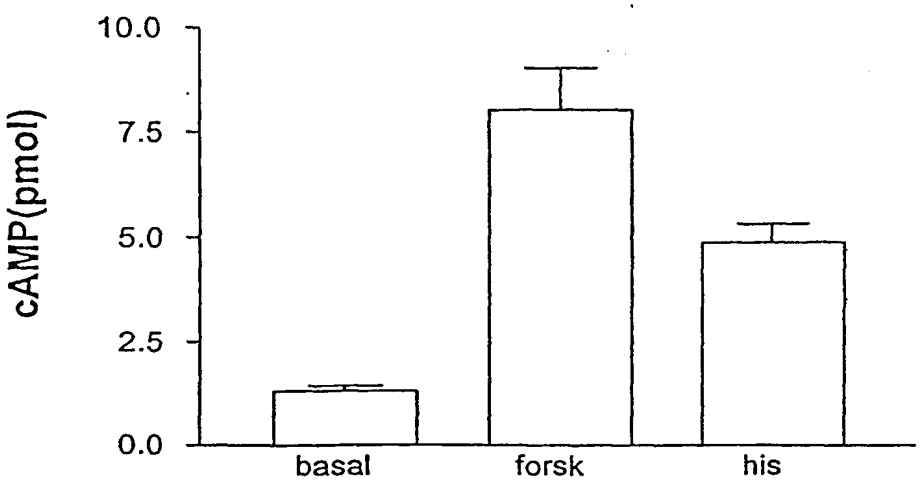


FIG. 3

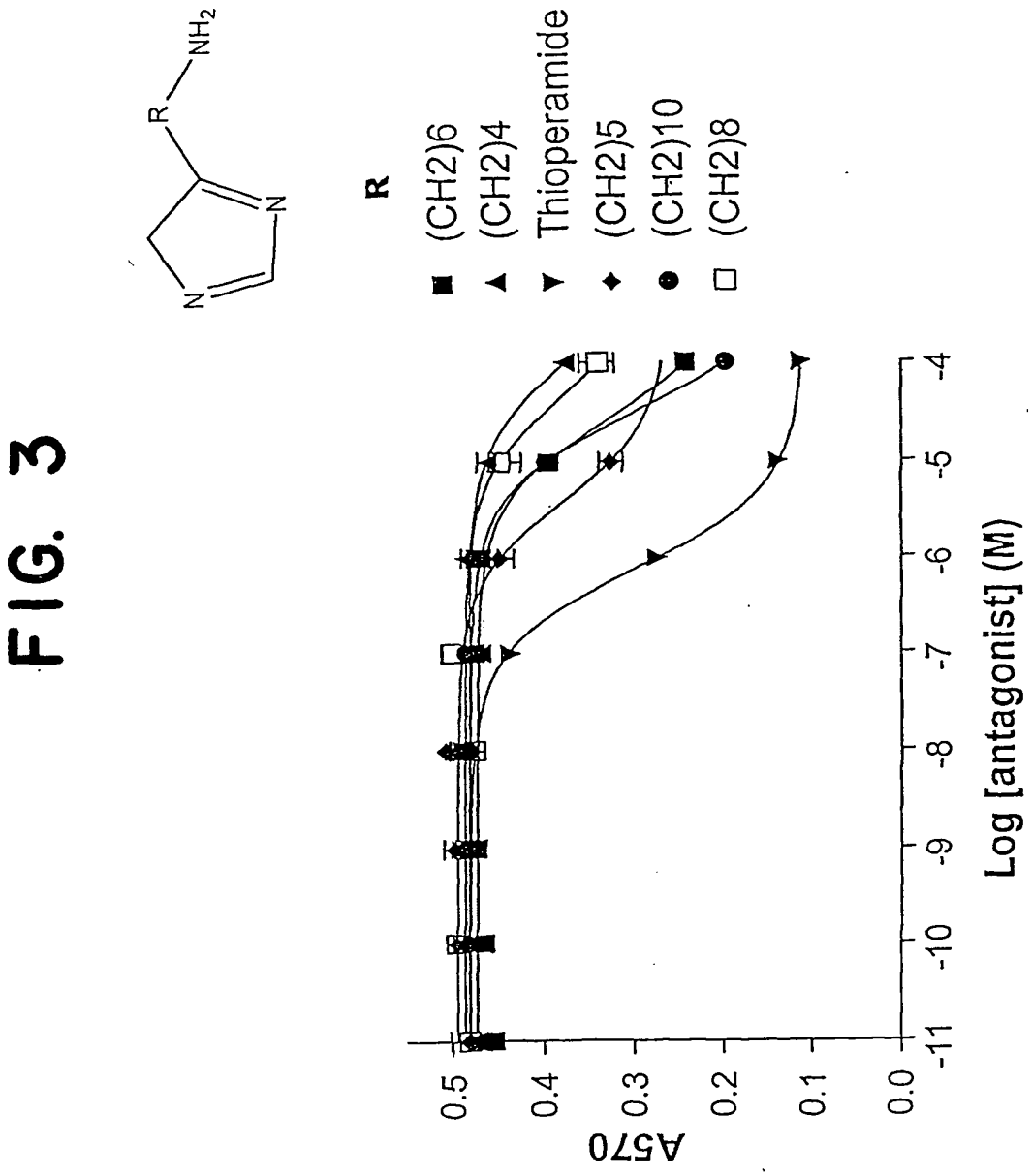

NH2
R
N
N
R
(CH2)6
(CH2)4
Thioperamide
(CH2)5
(CH2)10
(CH2)8
A570
0.5
0.4
0.3
0.2
0.1
0.0
-11
-10
-9
-8
-7
-6
-5
-4
Log [antagonist] (M)

FIG. 4

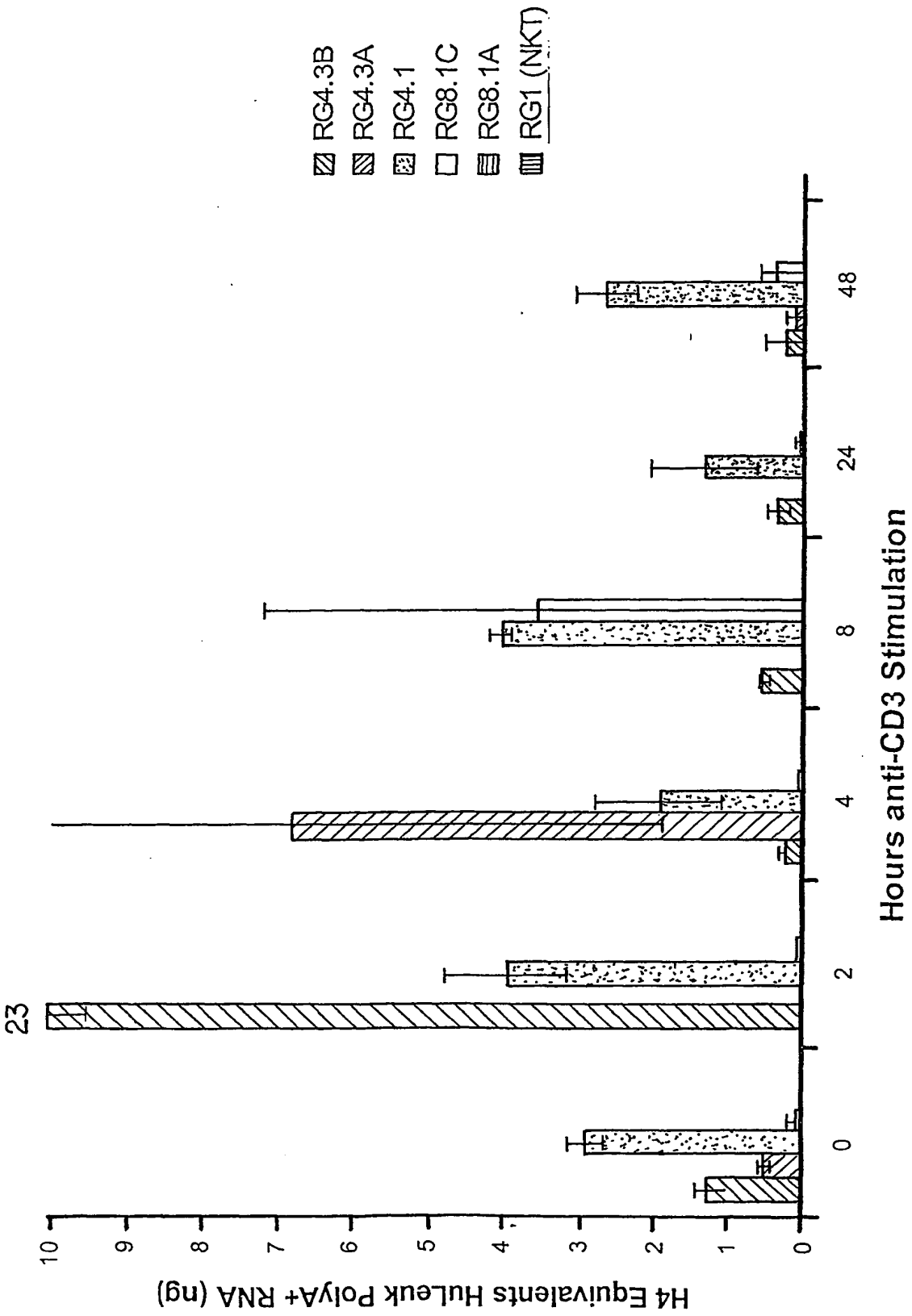
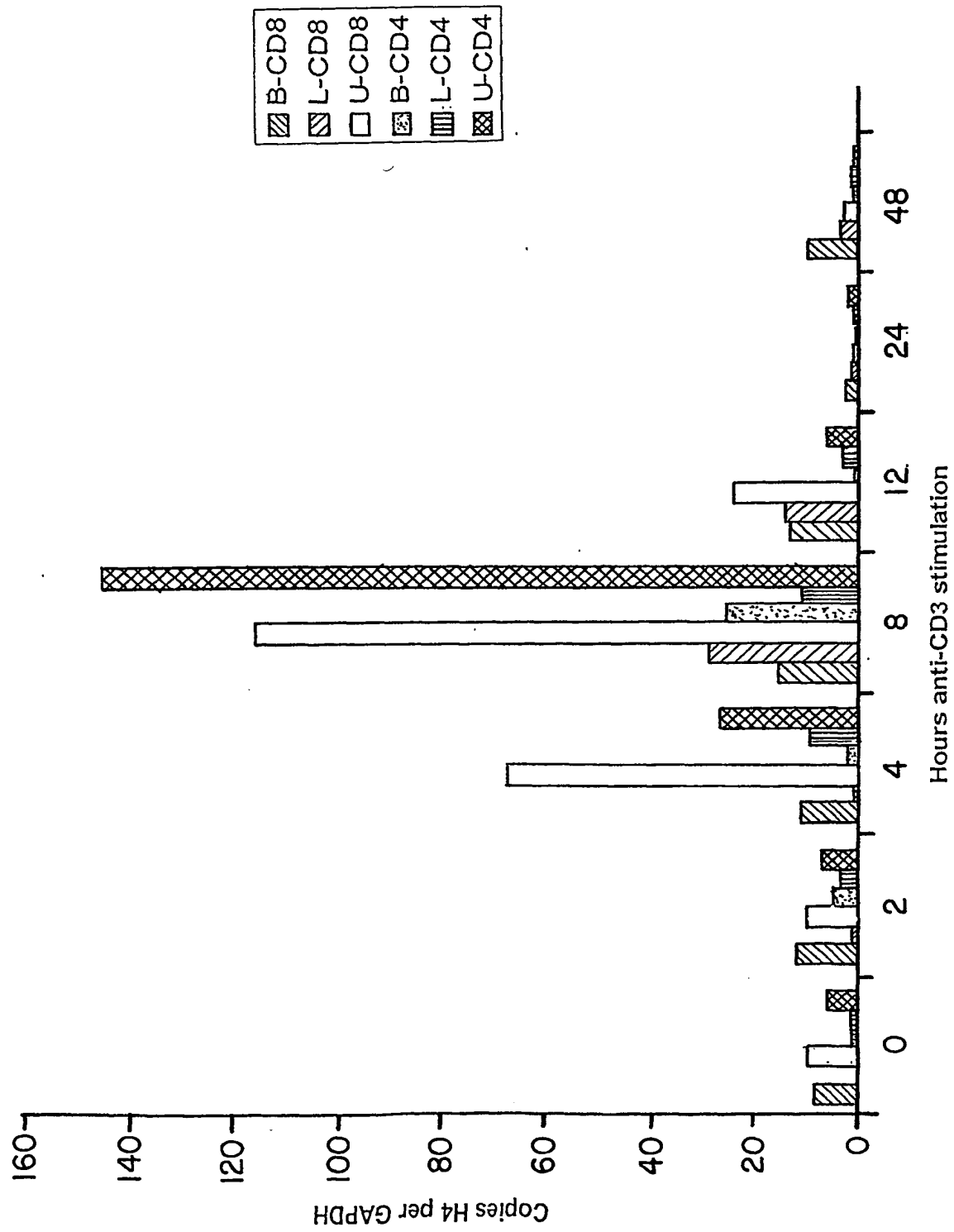

RG4.3B
RG4.3A
RG4.1
RG8.1C
RG8.1A
RG1 (NKT)
H4 Equivalents HuLeuk PolyA+ RNA (ng)
0
1
2
3
4
5
6
7
8
9
10
23
0
2
4
8
24
48
Hours anti-CD3 Stimulation

FIG. 5


B-CD8
L-CD8
U-CD8
B-CD4
L-CD4
U-CD4
Copies H4 per GAPDH
0
20
40
60
80
100
120
140
160
0
2
4
8
12
24
48
Hours anti-CD3 stimulation

SEQUENCE LISTING

<110> Jones, Philip
Shujian, Wu
Blatcher, Maria
Pausch, Mark

<120> Human Histamine H4 Receptor

<130> 3655/1J100-US3

<140> TBA
<141> Concurrently Herewith

<150> 60/202,151
<151> 2000-05-05

<150> 60/227,567
<151> 2000-08-23

<150> 60/247,855
<151> 2000-11-13

<160> 15

<170> PatentIn version 3.0

<210> 1
<211> 1173
<212> DNA
<213> Homo sapiens

<400> 1
atgccagata ctaatagcac aatcaattta tcactaagca ctcgtgttac tttagcattt 60

tttatgtcct tagtagcttt tgctataatg ctaggaaatg ctttggtcat tttagctttt 120

gtggtggaca aaaaccttag acatcgaagt agttattttt ttcttaactt ggccatctct 180

gacttctttg tgggtgtgat ctccattcct ttgtacatcc ctcacacgct gttcgaatgg 240

gattttggaa aggaaatctg tgtattttgg ctcactactg actatctgtt atgtacagca 300

tctgtatata acattgtcct catcagctat gatcgatacc tgtcagtctc aaatgctgta 360

agttatagaa ctcaacatac tggggtcttg aagattgtta ctctgatggt ggccgtttgg 420

gtgctggcct tcttagtgaa tgggccaatg attctagttt cagagtcttg gaaggatgaa 480

ggtagtgaat gtgaacctgg atttttttcg gaatggtaca tccttgccat cacatcattc 540

ttggaattcg tgatcccagt catcttagtc gcttatttca acatgaatat ttattggagc 600

ctgtggaagc gtgatcatct cagtaggtgc caaagccatc ctggactgac tgctgtctct 660

tccaacatct gtggacactc attcagaggt agactatctt caaggagatc tctttctgca 720

tcgacagaag ttcctgcatc ctttcattca gagagacaga ggagaaagag tagtctcatg 780

ttttcctcaa gaaccaagat gaatagcaat acaattgctt ccaaaatggg ttccttctcc 840

caatcagatt ctgtagctct tcaccaaagg gaacatgttg aactgctcag agccaggaga 900

ttagccaagt cactggccat tctcttaggg gtttttgctg tttgctgggc tccatattct 960

ctgttcacaa ttgtcctttc attttattcc tcagcaacag gtcctaaatc agtttggtat 1020

agaattgcat tttggcttca gtggttcaat tcctttgtca atcctctttt gtatccattg 1080

tgtcacaagc gctttcaaaa ggctttcttg aaaatatttt gtataaaaaa gcaacctcta 1140

ccatcacaac acagtcggtc agtatcttct taa 1173

<210> 2
<211> 390
<212> PRT
<213> Homo sapiens

<400> 2

Met Pro Asp Thr Asn Ser Thr Ile Asn Leu Ser Leu Ser Thr Arg Val
1 5 10 15

Thr Leu Ala Phe Phe Met Ser Leu Val Ala Phe Ala Ile Met Leu Gly
20 25 30

Asn Ala Leu Val Ile Leu Ala Phe Val Val Asp Lys Asn Leu Arg His
35 40 45

Arg Ser Ser Tyr Phe Phe Leu Asn Leu Ala Ile Ser Asp Phe Phe Val
50 55 60

Gly Val Ile Ser Ile Pro Leu Tyr Ile Pro His Thr Leu Phe Glu Trp
65 70 75 80

Asp Phe Gly Lys Glu Ile Cys Val Phe Trp Leu Thr Thr Asp Tyr Leu
85 90 95

Leu Cys Thr Ala Ser Val Tyr Asn Ile Val Leu Ile Ser Tyr Asp Arg
100 105 110

Tyr Leu Ser Val Ser Asn Ala Val Ser Tyr Arg Thr Gln His Thr Gly
115 120 125

Val Leu Lys Ile Val Thr Leu Met Val Ala Val Trp Val Leu Ala Phe
130 135 140

Leu Val Asn Gly Pro Met Ile Leu Val Ser Glu Ser Trp Lys Asp Glu

145 150 155 160

Gly Ser Glu Cys Glu Pro Gly Phe Phe Ser Glu Trp Tyr Ile Leu Ala
165 170 175

Ile Thr Ser Phe Leu Glu Phe Val Ile Pro Val Ile Leu Val Ala Tyr
180 185 190

Phe Asn Met Asn Ile Tyr Trp Ser Leu Trp Lys Arg Asp His Leu Ser
195 200 205

Arg Cys Gln Ser His Pro Gly Leu Thr Ala Val Ser Ser Asn Ile Cys
210 215 220

Gly His Ser Phe Arg Gly Arg Leu Ser Ser Arg Arg Ser Leu Ser Ala
225 230 235 240

Ser Thr Glu Val Pro Ala Ser Phe His Ser Glu Arg Gln Arg Arg Lys
245 250 255

Ser Ser Leu Met Phe Ser Ser Arg Thr Lys Met Asn Ser Asn Thr Ile
260 265 270

Ala Ser Lys Met Gly Ser Phe Ser Gln Ser Asp Ser Val Ala Leu His
275 280 285

Gln Arg Glu His Val Glu Leu Leu Arg Ala Arg Arg Leu Ala Lys Ser
290 295 300

Leu Ala Ile Leu Leu Gly Val Phe Ala Val Cys Trp Ala Pro Tyr Ser
305 310 315 320

Leu Phe Thr Ile Val Leu Ser Phe Tyr Ser Ser Ala Thr Gly Pro Lys
325 330 335

Ser Val Trp Tyr Arg Ile Ala Phe Trp Leu Gln Trp Phe Asn Ser Phe
340 345 350

Val Asn Pro Leu Leu Tyr Pro Leu Cys His Lys Arg Phe Gln Lys Ala
355 360 365

Phe Leu Lys Ile Phe Cys Ile Lys Lys Gln Pro Leu Pro Ser Gln His
370 375 380

Ser Arg Ser Val Ser Ser
385 390

<210> 3
<211> 259
<212> PRT
<213> Artificial

<220>
<223> Consensus sequence

<400> 3

Gly Asn Leu Leu Val Ile Leu Val Ile Leu Arg Thr Lys Lys Leu Arg
1 5 10 15

Thr Pro Thr Asn Ile Phe Ile Leu Asn Leu Ala Val Ala Asp Leu Leu
20 25 30

Phe Leu Leu Thr Leu Pro Pro Trp Ala Leu Tyr Tyr Leu Val Gly Gly
35 40 45

Ser Glu Asp Trp Pro Phe Gly Ser Ala Leu Cys Lys Leu Val Thr Ala
50 55 60

Leu Asp Val Val Asn Met Tyr Ala Ser Ile Leu Leu Leu Thr Ala Ile
65 70 75 80

Ser Ile Asp Arg Tyr Leu Ala Ile Val His Pro Leu Arg Tyr Arg Arg
85 90 95

Arg Arg Thr Ser Pro Arg Arg Ala Lys Val Val Ile Leu Leu Val Trp
100 105 110

Val Leu Ala Leu Leu Leu Ser Leu Pro Pro Leu Leu Phe Ser Trp Val
115 120 125

Lys Thr Val Glu Glu Gly Asn Gly Thr Leu Asn Val Asn Val Thr Val
130 135 140

Cys Leu Ile Asp Phe Pro Glu Glu Ser Thr Ala Ser Val Ser Thr Trp
145 150 155 160

Leu Arg Ser Tyr Val Leu Leu Ser Thr Leu Val Gly Phe Leu Leu Pro
165 170 175

Leu Leu Val Ile Leu Val Cys Tyr Thr Arg Ile Leu Arg Thr Leu Arg
180 185 190

Lys Ala Ala Lys Thr Leu Leu Val Val Val Val Val Phe Val Leu Cys
195 200 205

Trp Leu Pro Tyr Phe Ile Val Leu Leu Leu Asp Thr Leu Cys Leu Ser
210 215 220

Ile Ile Met Ser Ser Thr Cys Glu Leu Glu Arg Val Leu Pro Thr Ala
225 230 235 240

Leu Leu Val Thr Leu Trp Leu Ala Tyr Val Asn Ser Cys Leu Asn Pro
245 250 255

Ile Ile Tyr

<210> 4

<211> 21
<212> DNA
<213> Artificial

<220>
<223> Forward primer

<400> 4
taacttggcc attgacttct t 21

<210> 5
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Reverse Primer

<400> 5
attcgaacag catgtgaggg at 22

<210> 6
<211> 27
<212> DNA
<213> Artificial

<220>
<223> Probe

<220>
<221> modified_base
<222> (1)..(1)
<223> (6-carboxyfluorescein)-t

<220>
<221> modified_base
<222> (27)..(27)
<223> a- (6-carboxy-N,N,N'N'-tetramethylrhodamine)

<400> 6
tacaaaggaa tggagatcac acccaca 27

<210> 7
<211> 29
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 7
aaggatccaa aatgccagat actaatagc 29

<210> 8
<211> 28
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 8
aagtcgactt aagaagatac tgaccgac 28

<210> 9
<211> 43
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 9
aagcttccac catgccagat actaatagca caatcaattt atc 43

<210> 10
<211> 33
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 10
tctagattaa gaagatactg accgactgtg ttg 33

<210> 11
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 11
ggaaggatga aggtagtgaa tg 22

<210> 12
<211> 22

<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 12
cagaatctga ttgggaggaa gg 22

<210> 13
<211> 19
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 13
taacttggcc atctttgac 19

<210> 14
<211> 18
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 14
attcgaacag cgtgtgag 18

<210> 15
<211> 19
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<220>
<221> modified_base
<222> (1)..(1)
<223> 5'-(6-carboxyfluorescein)-t

<400> 15
tacaaaggaa tggagatca 19

INTERNATIONAL SEARCH REPORT

Inte onal Application No
PCT/US 01/14527

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C07K14/72 C12N15/16 C12N5/10 A01K67/027 C07K16/28
C12Q1/68 G01N33/566

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K C12N A01K C12Q G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, BIOSIS, SEQUENCE SEARCH

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 00 22131 A (ARENA PHARMACEUTICALS INC - GORE ET AL) 20 April 2000 (2000-04-20) * See SEQ.ID.NOS 14 (list pages 18-19) and Example 1 (pages 23-24) *	1-33
A	LOVENBERG T W ET AL: "Cloning and functional expression of the human histamine H3 receptor" MOLECULAR PHARMACOLOGY, BALTIMORE, MD, US, vol. 55, 1999, pages 1101-1107, XP002942531 ISSN: 0026-895X * See page 1103, Figure 1 (GPCR97) *	1-33

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 December 2001	02/01/2002

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Korsner, S-E

Form PCT/ISA/210 (second sheet) (July 1992)